1994

Three-Dimensional Radiative Transfer on a Massively Parallel Computer.

Horst Michael Vath
Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_disstheses/5839

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Three-dimensional radiative transfer on a massively parallel computer

Vath, Horst Michael, Ph.D.
The Louisiana State University and Agricultural and Mechanical Col., 1994
THREE-DIMENSIONAL RADIATIVE
TRANSFER ON A MASSIVELY PARALLEL
COMPUTER

A Dissertation

Submitted to the Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Physics and Astronomy

by
Horst M. Väth
M.S., Louisiana State University, 1990
August 1994
ACKNOWLEDGEMENTS

The years at LSU have been a great experience. For this I particularly have to thank Dr. D. Koester. He was my advisor for the first two years of my work towards the Ph.D., and those two years are generally known to be the most difficult ones. Thanks to his patience and his willingness for discussions I was able to make significant progress in those two years.

After Dr. Koester had left LSU, Dr. G. Chanmugam, Dr. J. Frank and Dr. J. Tohline acted as my effective advisors during the last year of my work for the Ph.D. thesis. It was a great pleasure to work with them and find astrophysical applications for my radiative transfer codes. Together with Dr. Tohline I worked on the structure of galaxies. With Dr. Chanmugam I investigated polarized radiative transfer and accretion columns onto magnetic white dwarfs. Dr. Frank also gave advise on this subject. I particularly want to point out the efforts of Dr. Chanmugam to help me in my work even while he was on a sabbatical in Boulder, Colorado. Here I also have to thank the other members of Dr. Tohline’s group, especially Kim Barker and Howard Cohl for their large contributions to the work on HI disks of galaxies.

Furthermore I want to thank all my committee members. Besides Dr. Chanmugam and Dr. Tohline these are Dr. Kalia, Dr. Perry, Dr. Metcalf and Dr. Wathugala as the outside committee member. In particular, I want to thank Dr. Kalia for his suggestions and comments on the computational aspects of this thesis.

I also want to thank Dr. Chanmugam for being the advisor for my Master’s thesis. Furthermore, I want to express my thanks to LSU and the
Ruprecht-Karls Universität of Heidelberg, Germany, as it was through the exchange program between these two universities that I initially came to LSU.

This work was supported in part by grants from the National Science Foundation (AST 90 8244) and NASA (NAGW 2447) and by the Louisiana State Board of Regents in providing funds to establish the Concurrent Computing Laboratory at the LSU Department of Physics and Astronomy.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ACKNOWLEDGEMENTS</th>
<th>ii</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>x</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2 THE MASPAR MP-1</td>
<td>7</td>
</tr>
<tr>
<td>3 IMAGING OF 3D ASTROPHYSICAL SYSTEMS</td>
<td>11</td>
</tr>
<tr>
<td>4 HI DISTRIBUTION IN WARPED GALAXIES</td>
<td>18</td>
</tr>
<tr>
<td>5 DUST IN OUTER REGIONS OF GALAXIES</td>
<td>34</td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>34</td>
</tr>
<tr>
<td>5.2 Galaxy model and radiative transfer</td>
<td>35</td>
</tr>
<tr>
<td>5.3 NGC 5055</td>
<td>40</td>
</tr>
<tr>
<td>5.4 Summary</td>
<td>49</td>
</tr>
<tr>
<td>6 POLARIZED RADIATIVE TRANSFER</td>
<td>51</td>
</tr>
<tr>
<td>6.1 Introduction</td>
<td>51</td>
</tr>
<tr>
<td>6.2 Cyclotron opacities</td>
<td>54</td>
</tr>
<tr>
<td>6.2.1 Introduction</td>
<td>54</td>
</tr>
<tr>
<td>6.2.2 Cyclotron opacities</td>
<td>57</td>
</tr>
<tr>
<td>6.2.3 Discussion</td>
<td>64</td>
</tr>
<tr>
<td>6.3 Accretion columns in AM Her binaries</td>
<td>68</td>
</tr>
<tr>
<td>6.4 3D effects in an extended polar cap</td>
<td>76</td>
</tr>
<tr>
<td>6.5 The intermediate polar RE 0751+14</td>
<td>81</td>
</tr>
<tr>
<td>7 NLTE RADIATIVE TRANSFER</td>
<td>89</td>
</tr>
<tr>
<td>7.1 Introduction</td>
<td>89</td>
</tr>
<tr>
<td>7.2 Radiative transfer</td>
<td>90</td>
</tr>
</tbody>
</table>
7.3 Solution of the NLTE radiative transfer problem
on the MasPar......................................................... 96
7.3.1 Short characteristics..................................... 96
7.3.2 Timing......................................................... 105
7.4 Discussion – Dirichlet boundary conditions............ 108
7.5 Discussion – Periodic boundary conditions.............. 117
7.6 Conclusion...................................................... 121

8 CONCLUSION.................................................. 125

REFERENCES....................................................... 129

APPENDICES
A POLARIZED RADIATIVE TRANSFER.................... 135
B LETTER OF PERMISSION..................................... 146
VITA................................................................. 147
<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>46</td>
</tr>
<tr>
<td>4</td>
<td>65</td>
</tr>
<tr>
<td>5</td>
<td>107</td>
</tr>
</tbody>
</table>

1. Timing for global router operation
2. Model parameters for the different galaxies
3. Dust in the outer regions of NGC 5055
4. Comparison between absorption coefficients $\alpha_{\pm}$ obtained using the Tamor method and using the Stokes method together with Eq. (6.20)
5. Scaling of the computing time for the different calculations on an isotropic cartesian grid of $N^3$ grid points
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A much simplified diagram of the main components of the MasPar MP-1</td>
</tr>
<tr>
<td>2</td>
<td>Geometry of input grid and rays used to solve the transfer equation</td>
</tr>
<tr>
<td>3</td>
<td>HI distribution of our model galaxy resembling NGC 2841</td>
</tr>
<tr>
<td>4</td>
<td>Same as Fig. 3 but for the galaxy NGC 7331</td>
</tr>
<tr>
<td>5</td>
<td>Same as Fig. 3 but for the galaxy M83</td>
</tr>
<tr>
<td>6</td>
<td>Same as Fig. 3 but for the galaxy NGC 5033</td>
</tr>
<tr>
<td>7</td>
<td>Same as Fig. 3 but for the galaxy NGC 5055</td>
</tr>
<tr>
<td>8</td>
<td>Same as Fig. 3 but for the galaxy NGC 300</td>
</tr>
<tr>
<td>9</td>
<td>Observational radial velocity map of NGC 5033</td>
</tr>
<tr>
<td>10</td>
<td>Observational column density map of NGC 5033</td>
</tr>
<tr>
<td>11</td>
<td>Observational radial velocity map of NGC 5055</td>
</tr>
<tr>
<td>12</td>
<td>Observational column density map of NGC 5055</td>
</tr>
<tr>
<td>13</td>
<td>Calculated HI map of NGC 5055</td>
</tr>
<tr>
<td>14</td>
<td>Three-dimensional rendering of the HI distribution</td>
</tr>
<tr>
<td>15</td>
<td>Optical image of our model galaxy</td>
</tr>
<tr>
<td>16</td>
<td>Contour plot of a CCD image of NGC 5055 in the B band</td>
</tr>
<tr>
<td>17</td>
<td>Same as Fig. 16 but this image of NGC 5055 is in the I band</td>
</tr>
<tr>
<td>18</td>
<td>Comparison of the opacities $\kappa$, $q$, $v$ calculated from the</td>
</tr>
</tbody>
</table>
Robinson & Melrose (1984) analytic formulae and the accurate numerical opacities for $kT = 10$ keV and different angles $\theta$ ................................................. 69

19 Same as Fig. 18 but for $\theta = 90^\circ$ and different temperatures ................................................. 70

20 Theoretical spectrum of the cyclotron emission of an AM Her system ............................................. 74

21 Geometry of a polar cap ................................................. 78

22 Cyclotron emission from an extended polar cap ........ 79

23 Rays emitted from an extended 3-dimensional polar cap 81

24 Images of the emission region of an extended polar cap 82

25 The model used to calculate the light curve of RE 0751+14 ......................................................... 84

26 Modeled light curve of RE 0751+14 over one period . . 86

27 Fit of the model to the observed maximum circularly polarized flux ............................................. 88

28 x-y layer of the 3D cartesian grid showing a short characteristic for point p .................................. 98

29 Number of steps needed for convergence to get specific intensities in one x-y layer .......................... 102

30 Solution and initial guess for the source function, diagonal of the $\Lambda$-matrix and optical depth $\tau$ ....... 109

31 Convergence of the classical $\Lambda$-iteration, ALI with local $\Lambda^*$ and ALI with non-local $\Lambda^*$ .................. 112

32 Maximum relative error of the source function after each iteration step ............................................ 114

33 Convergence of Ng acceleration, orthomin with local $\Lambda^*$ and orthomin with non-local $\Lambda^*$ ............ 116

34 Convergence of multi-grid methods with $n = 2, n = 3,$
and \( n = 4 \) .......................................................... 118

35 Part of grey 3D atmosphere with periodic boundary
conditions along \( x \) and \( y \) .............................................. 119

36 Maximum relative error of the source function after each
iteration step ................................................................. 122

37 Calculated value and value in case of large Faraday
rotation and Faraday pulsation of the ratio \( Q/V \) for
\( kT = 20 \) keV ............................................................ 143

38 Same as the previous figure but for \( kT = 50 \) keV ...... 144
ABSTRACT

We perform three-dimensional radiative transfer calculations on the MasPar MP-1, which contains 8192 processors and is a single instruction multiple data (SIMD) machine, an example of the new generation of massively parallel computers. To make radiative transfer calculations efficient, we must re-consider the numerical methods and methods of storage of data that have been used with serial machines. We developed a numerical code which efficiently calculates images and spectra of astrophysical systems as seen from different viewing directions and at different wavelengths. We use this code to examine a number of different astrophysical systems.

First we image the HI distribution of model galaxies. Then we investigate the galaxy NGC 5055, which displays a radial asymmetry in its optical appearance. This can be explained by the presence of dust in the outer HI disk far beyond the optical disk. As the formation of dust is connected to the presence of stars, the existence of dust in outer regions of this galaxy could have consequences for star formation at a time when this galaxy was just forming.

Next we use the code for polarized radiative transfer. We first discuss the numerical computation of the required cyclotron opacities and use them to calculate spectra of AM Her systems, binaries containing accreting magnetic white dwarfs. Then we obtain spectra of an extended polar cap. Previous calculations did not consider the three-dimensional extension of the shock. We find that this results in a significant underestimate of the radiation emitted in the shock. Next we calculate the spectrum of the intermediate polar RE
0751+14. For this system we obtain a magnetic field of \( \sim 10 \) MG, which has consequences for the evolution of intermediate polars.

Finally we perform 3D radiative transfer in NLTE in the two-level atom approximation. To solve the transfer equation in this case, we adapt the short characteristic method and examine different acceleration methods to obtain the source function. These include the ALI method with local and non-local operators, the Ng and the orthomin methods and multi-grid methods. We apply these numerical methods to two problems with and without periodic boundary conditions.
CHAPTER 1

INTRODUCTION

In astronomy one generally observes the radiation emitted from distant systems. In order to understand the physical structure of such an astronomical system one has to know how the system produces the radiation observed at a given wavelength. The transfer of radiation through matter is described by the radiative transfer equation. One therefore has to be able to solve this transfer equation for a model of the observed system. One can then compare the calculated spectrum of the model with the observations, and by finding the model that fits the observations best, one can draw conclusions on the physical structure of the system.

The classical radiative transfer problem in astronomy has been concerned with calculating spectra of stellar atmospheres at rest (Mihalas 1978). Because of spherical symmetry, this is essentially a one-dimensional problem with the radius as the only spatial coordinate. Several powerful methods for solving the radiative transfer equation for this case have been known for many years (Feautrier 1964; Auer & Mihalas 1969; Rybicki 1971), and it has become possible to include in such calculations many line transitions. In addition, the statistical and radiative equilibrium equations and the hydrostatic equations can be solved self-consistently (e.g. Werner 1989; Dreizler 1992). In three dimensions, however, it has not been possible to do this on present computers.
so far, both because the size of the problem exhausts the memory space, and because such computations would require far too much computing time. Three-dimensional calculations become necessary for astrophysical systems that are not spherically symmetric. Accretion disks are such systems (Adam 1990). Other examples are certain planetary nebulae (Icke et al. 1992) and gas clouds generated by collisions between stars (Ruffert 1992). To calculate spectra of such systems, one therefore needs to solve the radiative transfer equation in three dimensions. The increasing power of computers has now made it possible to tackle such problems. Several approaches to solving the transfer equations in more than one dimension have been made in recent years. Stenholm et al. (1991) computed the specific intensity at every point of a 3D cartesian grid by approximating the transfer equation as a difference equation. Another way of solving the transfer equation is by using the so-called short characteristics (Kunasz & Auer 1988). Recently Klein et al. (1989) applied the finite element method to this problem. Except for the finite difference method, all these methods have only been applied to 2D cases so far.

In the case of non-local thermal equilibrium (NLTE), the source function, on which the radiation field depends, is itself a function of the radiation field. The transfer equation then becomes an integral-differential equation and the calculation of the source function is very difficult. A much easier situation arises when matter is in local thermal equilibrium (LTE). Then the source function can be calculated directly from local quantities such as the temperature and the density. But whether one has obtained the source function from an NLTE or simply from an LTE calculation, one is still left with the task of
calculating the image of the 3D system as seen from different directions and at different wavelengths. Only then can one compare the model with the actual observations. As the image of an astrophysical system can vary strongly with viewing direction and wavelength, one needs a tool which solves the 3D radiative transfer problem rapidly.

Traditionally, all radiative transfer calculations have been performed on scalar and vector computers. However, in recent years massively parallel computers have become available in a variety of different architectures. One way to classify these computers is by the way the machine relates the instructions to the data (Flynn 1972; Hockney & Jesshope 1988). Then two major classes of parallel machines can be distinguished. On the one hand, there are the multiple instruction multiple data (MIMD) machines, on which each processor can make computations completely independently of other processors. But the individual processors of such computers have to be fairly complex, are therefore relatively expensive, and the number of processors is usually small. The advantage of such computers is that a program can be run on different processors with different parameters without making adjustments within the program. In contrast to these MIMD computers are the single instruction multiple data (SIMD) machines, on which all processors perform exactly the same computations. The processors therefore can be relatively simple and are inexpensive, and such computers can contain a large number of processors. There the code mainly consists of array manipulations, and a language in which parallel array operations are easily defined is required to program such computers.
efficiently. Overall, it cannot be decided yet which approach in parallel computers will prevail, and currently applications are designed for both types of computers, which is reflected in conference proceedings on parallel computing (e.g. Meuer 1991).

In this thesis we discuss various aspects of 3D radiative transfer. All these calculations have been performed on the SIMD computer MasPar MP-1. To our knowledge, no radiative transfer calculations have been performed so far on SIMD machines. First we describe the MasPar MP-1. In the following chapter we describe a numerical tool, that we have developed, to view the model of an astrophysical system rapidly from different directions and at different wavelengths. We then use this tool in order to investigate various astrophysical systems. First we study the distribution of neutral hydrogen (HI) in galaxies as its radiation is the main tracer for the large scale structure of disk galaxies. We calculated the HI profile of our model galaxy as it appears from different viewing directions and have thus produced a movie of a "fly-by" around our model galaxy. Here we display some images from this video in which the calculated profiles resemble those of observed galaxies. In Chapter 5 we examine observations of the galaxy NGC 5055 both at optical wavelengths and in the 21 cm line radiation of neutral hydrogen. We suggest that dust is present in the HI disk at radii far beyond the optical disk. The presence of dust at such radii could have important consequences for the evolution of disk galaxies as the formation of dust is generally linked to the presence of stars. In order to test this hypothesis we use our imaging tool both in the 21
cm line of HI and in the optical wavelength regions in order to reproduce the observations.

A third area that we investigate with the imaging tool, which is contained in Chapter 6, involves radiative transfer in accretion columns onto magnetic white dwarfs. Because of the presence of a strong magnetic field one must solve the transfer equation for polarized radiation. If one can model the light emitted from such systems one can obtain the strength of the magnetic field. This in turn can be important for understanding the physical structure and the evolution of these systems. In Chapter 6 we start with a description of the recent work done in this field and how our work fits into it. Then we present a critical discussion of the numerical calculation of the necessary opacities. Such a discussion does not exist in the literature for these opacities. We then use the opacities to calculate the spectra of AM Her systems which are binaries containing accreting white dwarfs. These systems show strong linear and circular polarization which implies strong magnetic fields (10 to 70 MG). In contrast the spectra of so-called intermediate polars (IP) are unpolarized. However, it has been suggested that these systems also have magnetic fields comparable to AM Her systems (Chanmugam & Ray 1984; King et al. 1985), and the absence of polarization may only be due to dilution effects. It was suggested that accretion onto a large surface area may dilute the polarization (Chanmugam & Frank 1987). This led us to investigate 3D effects in extended polar caps. Finally we apply our code to model the IP RE 0751 +14 discovered recently by the ROSAT X-ray satellite. This system is important because it
shows both linear and circular polarization (Pirola et al. 1993) and may therefore represent a link between AM Her systems and IPs. In a fourth application we have used our imaging tool to calculate combined line and continuum spectra of accretion disk from the ultraviolet to the near infrared (Väth 1994a). However, we will not discuss this here.

In all astrophysical systems discussed above the matter is in LTE. However, there are also many systems in which the matter is in NLTE. Therefore we discuss in Chapter 7 NLTE radiative transfer in three-dimensional mass distributions. In NLTE radiative transfer one must obtain the source function iteratively, which is computationally very expensive. We test various numerical methods in order to accelerate the convergence. A detailed description of this part of the thesis is also given in Väth (1994b). Final conclusions of this thesis are presented in Chapter 8.
CHAPTER 2
THE MASPAR MP-1

When programming on conventional computers, one needs to know very little to nothing at all about the architecture of the computer. This is not the case for massively parallel computers. There the architecture directly influences the style of computer codes, as one has to think in terms of how data are flowing macroscopically between different microprocessors when the program is executed. Even further, the existence of a certain architecture might make the solution of one problem very easy and fast, while essentially prohibiting the treatment of another problem on this particular computer. Because of these reasons, we have to give a fairly detailed description of the architecture of the MP-1. However, we mostly restrict ourselves to those features of the MP-1 that are important to us here. (The following description is based on the manuals available for the MasPar MP-1.)

The MP-1 is a SIMD machine, which means that all processors execute the same commands. It contains 8192 processing elements (PEs), although this number can be extended to 16384. The basic components of the MP-1 are shown in Fig. 1. The PEs are arranged in a two-dimensional grid. The data memory of the PEs then forms a natural z direction. On the MP-1 at the Dept. of Physics & Astronomy, Louisiana State University, on which all calculations have been performed, there are 128 PEs in the x and 64 PEs in the y direction.
Figure 1. A much simplified diagram of the main components of the MasPar MP-1. The PE-grid is a two-dimensional array of processors each with its own memory. All PEs perform the same operation but on different data. These operations are orchestrated by the ACU, which can also communicate with the FE. The global router allows transfer of data between different sets of PEs concurrently.

Each processor contains 64 kByte of memory space. As all processors perform the same operations, they need to be controlled by some central unit, which is the array control unit (ACU). While all parallel operations are performed on the PE grid, scalar calculations are done either on the ACU or the front end (FE), which is here a DEC 5000 workstation. Communication between the FE and the PE grid is possible through the ACU, which can broadcast scalar data simultaneously onto all PEs.

An ideal problem for such a SIMD machine would be one, in which all the processors could operate independently from one another, i.e. no data
would need to be transferred between processors. However, in most realistic applications, this is not the case. We therefore have to describe the way processors can communicate with each other. On the MasPar there are essentially two ways to do that. Each processor on the PE grid is directly connected with the surrounding eight processors. This so-called X-NET enables the transfer of entire floating point numbers, and communication is therefore very fast. But this method of data transfer is not useful, if two processors which have to communicate with each other are very far apart on the PE grid. For this case the global router can be used, which can connect any two PEs concurrently. Up to 512 simultaneous communications are possible on the MP-1. Because of the smaller number of possible connections between PEs, this type of communication is relatively slow.

To program a SIMD machine, one needs a computer language that allows parallel manipulation of data in a simple way. On the MP-1, two languages are currently available: MPL and MPFORTRAN. The programming language MPL is an enhancement of C and is the fundamental programming language on the MP-1. The code described here, however, is written in MPFORTRAN, which is close to the new FORTRAN 90 standard and is very suitable for parallel machines because of its array syntax. It also has the advantage that the programmer does not need to explicitly consider the number of available processors (although this number strongly influences the execution time of the program). If one has for example a three-dimensional array of some arbitrary size, then the first two dimensions are automatically allocated along the $x$ and $y$ axis of the PE grid, while the third dimension is stored in the
memory of the PEs. If the $x$ and $y$ dimensions of the array are too large to fit onto the PE grid, the array is cut and stacked into the memory of the PEs and forms virtual layers there. (It is possible, but normally not necessary, to override this default.) Furthermore, the programmer does not need to decide explicitly how communications have to be performed, as is necessary in MPL, which makes the programming less tedious. It has the additional advantage that the code is less specialized for the architecture of the MP-1 and therefore can be adapted to different SIMD machines or conventional scalar machines more easily once the FORTRAN 90 standard is in general use.
CHAPTER 3

IMAGING OF 3D ASTROPHYSICAL SYSTEMS

To calculate the spectrum of a system in three dimensions, one must solve the radiative transfer equation (Mihalas 1978)

\[(\mathbf{n} \cdot \nabla) I_{\nu}(r, \mathbf{n}) = (d/ds)I_{\nu}(r, \mathbf{n}) = \chi_{\nu}(r, \mathbf{n})(S_{\nu}(r, \mathbf{n}) - I_{\nu}(r, \mathbf{n})) \tag{3.1}\]

where \(\chi\) is the opacity, \(S\) is the source function, \(I\) is the specific intensity, and no time dependence is considered. The vector \(\mathbf{n}\) gives the direction of the ray, \(\nu\) is the frequency, and \(r\) gives the position. The parameter \(s\) is the length along the ray propagating in the direction \(\mathbf{n}\). The transfer equation can be integrated easily along the line of sight, and thus one obtains the formal solution

\[I_{\nu}(s) = I_{\nu}(0) \exp(-\tau_{\nu}(s)) + \int_{0}^{\tau_{\nu}(s)} d\tau'_{\nu} S_{\nu}(\tau'_{\nu}) \exp(\tau'_{\nu} - \tau_{\nu}(s)) \tag{3.2}\]

\[\tau_{\nu}(s) = \int_{0}^{s} \chi_{\nu}(s')ds' \tag{3.3}\]

where \(\tau\) is the optical depth along the line of sight. In stellar atmospheres the radius is normally the only spatial variable, and one needs to consider only one angle for the vector \(\mathbf{n}\). However, if the system has no spherical symmetry, all quantities depend in general on three spatial variables, and \(\mathbf{n}\) is a function of two angles. As the opacity is normally also frequency dependent, the specific intensity becomes a function of six variables.
In order to image the model of an astrophysical system, one simply has to solve Eq. (3.1) for many rays. If these rays form a regular square when viewed by the observer, then the calculated image can be compared directly with images made with charge coupled devices (CCDs) of the real astronomical system. In Fig. 2 we show this situation. The input data are given on a cartesian grid in the unprimed coordinate system with $N_x \times N_y \times N_z$ grid points. On the MasPar MP-1 the input data are naturally mapped onto the PE grid (with dimensions $N_{xPE} \times N_{yPE}$ and a total number of PEs $N_{PE} = N_{xPE} \times N_{yPE}$) The x and y axis of the input grid is aligned with the x and y axis of the PE grid. The z axis of the input grid is stored in the memory of the PEs (if $N_x > N_{xPE}$ and/or $N_y > N_{yPE}$ then the arrays that stores the data of the input data are cut and stacked as described in the previous chapter). Each node of the input grid (and the PE that corresponds to this node) carries all the important information of the condition of matter at that point, e.g. the density and the temperature in LTE or the occupation numbers of atomic levels in NLTE. We denote the number of physical variables on each node by $N_{var}$.

In Fig. 2 we show some rays that pass through the input grid. The direction vector $\mathbf{n}$ of these rays defines a $z'$ axis. The rays form a regular grid with $N_{ray}$ rays in the $x'$ and in the $y'$ direction. One way to describe the (primed) coordinate system of the rays with respect to the (unprimed) coordinate system fixed in the input grid is by using the Euler angles $\Theta$ and $\Phi$ (Goldstein 1980). These two angles completely determine the components of the direction vector $\mathbf{n}$, which has the coordinates (0,0,1) in the system of
Figure 2. Geometry of input grid and rays used to solve the transfer equation. The cartesian input grid determines the unprimed coordinate system while the rays determine the primed coordinate system.

The observer. A third angle $\Psi$ can be used to define the $x'$ and $y'$ axis in the observer's frame. This last angle describes a rotation around the $z'$ axis and it therefore corresponds only to the position angle of the system in the sky. To go from the frame of some astronomical system into the frame of the observer, we use three successive rotations expressed by the transformation matrix (Goldstein 1980)

$$A = \begin{pmatrix} \cos \Psi & \sin \Psi & 0 \\ -\sin \Psi & \cos \Psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \Theta & \sin \Theta \\ 0 & -\sin \Theta & \cos \Theta \end{pmatrix} \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$  \hfill (3.4)

The direction vector $n$ in the frame of the astronomical system in terms of the Euler angles is $n = (\sin \Theta \sin \phi, -\sin \Theta \cos \phi, \cos \Theta)$. This vector transforms into the vector $n' = (0, 0, 1) = An$ in the observer's frame.
When solving the transfer equation (3.1) one naturally identifies each PE with a different ray. As the rays form a regular cartesian grid of $N_{\text{ray}}$ points in the $x'$ and $y'$ dimension of the observer's frame it is clear how the rays can best be identified with the PEs. Each ray has $N_b$ points, which are equidistant from each other. Once we have all relevant data on these $N_b$ points for some ray on the PE corresponding to that ray, we can use the data to compute the absorption coefficient and the source function. After that we can solve the transfer equation by calculating the integrals in the Eqs. (3.2) and (3.3). Clearly, all important information is already stored on the PEs which need it, and the calculations can be done for all rays simultaneously. The MasPar is therefore ideal for this, especially when the transfer equation has to be solved at many different wavelengths for the same viewing direction. However, to use the MasPar this way, we first have to transfer the data from the input grid onto the grid formed by the rays. Clearly, this will in general involve a macroscopic transfer of information between different PEs. As the information needed by some PE may be stored on any other PE, the best way to transfer the data is by using the global router.

In the following, we want to give some examples for the timing for different parts of the program. As stated above the input grid consists of $N_x \times N_y \times N_z$ points. At each point we have $N_{\text{var}}$ physical variables necessary to describe the state of the matter. We use $N_{\text{ray}}^2$ rays ($N_{\text{ray}}$ in both the $x$ and $y$ dimension in the observer's frame) and we have $N_b$ points along each ray. Here $N_{\text{ray}} = 128$. On the MasPar available to us we have $N_{x\text{PE}} = 128$ and $N_{y\text{PE}} = 64$. The total number of PEs is then $N_{\text{PE}} = 8192$. We therefore have
to divide the rays into two sets (each with $N_{xPE} \times N_{yPE}$ rays) and perform the entire calculations twice. That means that for one viewing direction we have to transfer the data from the input grid to the rays twice and we have to solve the transfer equation twice. On a machine with 128x128 PEs we will therefore only need half the time to calculate the image of a system.

A large amount of computing time is needed for the transfer of the data from the input grid (and the PEs corresponding to it) to the grid formed by the rays and their corresponding PEs. This time depends in a complicated way on the vector $n$, the size of the input grid and the size of the grid onto which the data is transferred (the latter is determined by $N_{ray}$ and $N_b$), as all this determines which PEs have to exchange information. While the timing is proportional to $N_{var}$, it is only approximately proportional to $N_b$. In Table 1 we list the timing for the transfer for different viewing angles and input grids but always use $N_{ray} = 128$ and $N_b = 256$. This table also lists the number of global router operations required for each transfer. As one can see from the table the transfer actually takes more time on a 64$^3$ input grid than on a 128$^3$ grid. On average it takes about 3 sec to transfer all the data corresponding to one variable. Only a few cases require more than 5 sec for this communication, while there also exists one case when the transfer only takes 0.25 sec. This occurs when the input grid is perfectly aligned with the grid formed by the rays. In this case no data have to be transferred from one PE to another and correspondingly the number of router operations is zero. In order to save communication time no data is transferred to a point on a ray that is outside the computational domain formed by the input grid. Once all the data have
been transferred, we need to calculate the specific intensity. For this part of the program the timing is proportional to $N_b$ and is independent of the number of rays. An exception occurs when there are more rays than PEs in which case the timing increases by a factor calculated by rounding down the number $1 + \left\lfloor \frac{N_{ray} - 1}{N_{PE}} \right\rfloor$ to its nearest integer. When for example the transfer equation is solved for neutral hydrogen in galaxies (see Chapter 5) on a MasPar with $128 \times 64$ PEs the calculation of the source function and opacity for HI takes 0.37 seconds for $N_b = 256$ and $128^2$ rays, and the calculation of the formal solution takes 0.17 seconds for each wavelength. If the calculations were done on a MasPar with $128 \times 128$ PEs the calculations would automatically take only half that time.

The above discussion has to be compared to a computer with only one processor. There the transfer of data from the input grid onto the rays scales as $N_{ray}^2 \times N_b \times N_{var}$ and on a scalar DEC 5000 workstation with $N_{ray} = 128$, $N_b = 128$, $N_{var} = 1$ and an input grid with $N = 64$ points per dimension the transfer took about 44 seconds (it is largely independent of the direction vector $n$). Once the transfer has been completed, all the other calculations scale as $N_{ray}^2 \times N_b$. 

Table 1. Timing for global router operation. Input data are stored on a regular cartesian grid of $N^3$ points with one physical variable per point. The data are transferred onto $128^3$ rays with 256 points on each ray on a MasPar with $128 \times 64$ PEs. The direction vector of the rays is characterized by the Euler angles $\Phi$ and $\Theta$ with $\Psi = 0$. The number of router operations and the time in milliseconds are given.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\Phi$</th>
<th>$\Theta$</th>
<th># of routers</th>
<th>Transfer time</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>0</td>
<td>0</td>
<td>158976</td>
<td>5721</td>
</tr>
<tr>
<td></td>
<td>$\pi/4$</td>
<td></td>
<td>69597</td>
<td>2647</td>
</tr>
<tr>
<td></td>
<td>$\pi/2$</td>
<td></td>
<td>100002</td>
<td>3684</td>
</tr>
<tr>
<td></td>
<td>$\pi/4$</td>
<td>0</td>
<td>98304</td>
<td>3631</td>
</tr>
<tr>
<td></td>
<td>$\pi/4$</td>
<td></td>
<td>45453</td>
<td>1824</td>
</tr>
<tr>
<td></td>
<td>$\pi/2$</td>
<td></td>
<td>70983</td>
<td>2697</td>
</tr>
<tr>
<td></td>
<td>$\pi/2$</td>
<td>0</td>
<td>147456</td>
<td>5327</td>
</tr>
<tr>
<td></td>
<td>$\pi/4$</td>
<td></td>
<td>75036</td>
<td>2834</td>
</tr>
<tr>
<td></td>
<td>$\pi/2$</td>
<td></td>
<td>85902</td>
<td>3196</td>
</tr>
<tr>
<td>128</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>246</td>
</tr>
<tr>
<td></td>
<td>$\pi/4$</td>
<td></td>
<td>43380</td>
<td>1742</td>
</tr>
<tr>
<td></td>
<td>$\pi/2$</td>
<td></td>
<td>81528</td>
<td>3055</td>
</tr>
<tr>
<td></td>
<td>$\pi/4$</td>
<td>0</td>
<td>55322</td>
<td>2167</td>
</tr>
<tr>
<td></td>
<td>$\pi/4$</td>
<td></td>
<td>31203</td>
<td>1320</td>
</tr>
<tr>
<td></td>
<td>$\pi/2$</td>
<td></td>
<td>77373</td>
<td>2918</td>
</tr>
<tr>
<td></td>
<td>$\pi/2$</td>
<td>0</td>
<td>98304</td>
<td>3642</td>
</tr>
<tr>
<td></td>
<td>$\pi/4$</td>
<td></td>
<td>44790</td>
<td>1802</td>
</tr>
<tr>
<td></td>
<td>$\pi/2$</td>
<td></td>
<td>75528</td>
<td>2845</td>
</tr>
</tbody>
</table>
CHAPTER 4

HI DISTRIBUTION IN WARPED GALAXIES

The 21 cm radiation of neutral hydrogen (HI) is an almost ideal method to measure the structure of galaxies because of the abundance of HI in the interstellar matter and because this radiation is normally optically thin. One can therefore use such observations to obtain a profile of the column density of a galaxy. This is done by observing the line profile of the 21 cm line along a certain line of sight and then integrating the intensity over the frequency. The column density of neutral hydrogen is (Mihalas & Binney 1981)

\[ N_H = 3.88 \times 10^{14} \int_0^\infty T_B d\nu \text{ cm}^{-2}. \]  

(4.1)

where the frequency units are in hertz and the brightness temperature \( T_B \) is defined as

\[ T_B \equiv \left( \frac{c^2}{2k\nu^2} \right) I_\nu. \]  

(4.2)

Furthermore, one can get from the measured line profile along a line of sight the radial velocity \( v_r \) where the density is maximal along the line of sight. If \( \nu_{\text{max}} \) is the frequency of the maximum intensity then

\[ v_r = c(\nu_0 - \nu_{\text{max}})/\nu_{\text{max}} \]  

(4.3)

where \( \nu_0 \) is the frequency of neutral hydrogen in the rest frame. Thereby, one can get a map of the column density and the radial velocity profile of the galaxy.
as seen by the observer. What one would like to have is a 3D model of the galaxy that can reproduce unambiguously the observed profiles. This can be done by constructing many 3D models of the galaxy, calculating the resulting column density and radial velocity profiles of each model and comparing these theoretical profiles with the observed ones. By using this procedure one can obtain the model which best fits the observations. This has been done by a number of people (Bosma 1981; Christodoulou et al. 1993, hereafter CTS). However, with our tool of calculating the image of a 3D astrophysical object it is particularly easy and fast to perform this procedure. In the following we describe the main features of the HI distribution in galaxies. After that we give the parameters typically used in models of galaxies. We thereby use the terminology as given by CTS. Finally, we image six galaxies with our code.

Neutral hydrogen gas normally rotates around the center of the galaxy in circular orbits. Any non-circular motion would be circularized after a few rotation periods, which are of the order of a few hundred million years, because of collisions between atoms and because the period is much shorter than the Hubble time \( H_0^{-1} \approx 10^{10} \) yrs except at extremely large distances from the center. From the observed distribution of HI gas, which can be detected at much larger distances from the center than the optical disk, one also comes to the conclusion that the matter is not situated in a flat disk but that this disk is warped along some axis and that this axis changes depending on the radius, i.e. the disk is both warped and twisted. This so-called tilted ring model of an HI disk was first used by Rogstad et al. (1974). Furthermore, Rubin et al. (1978) discovered that the rotation velocity of the matter outside the central
bulge (which is of no concern here) is approximately constant in most spiral galaxies. This is contrary to what one would expect from the observed mass distribution, which would result in a decrease of the velocity with the radius \( R \) as approximately \( R^{-1/2} \) from Kepler’s law. In order to explain this, one can either assume that Newton’s gravitational law is no longer valid over distances of several kiloparsec (Tohline 1983; Milgrom, 1983a, b, c; review by Sanders 1990) or, more commonly, one can assume the presence of dark matter, i.e. matter that cannot be observed. Typical models of this dark matter attribute to it a spheroidal potential with a quadrupole distortion, and two cases can be distinguished. In the first the potential has a pancake shape and is said to be oblate. In the second case, the potential has a cigar-like shape and is said to be prolate. When differentially rotating gas settles in a non-spherical gravitational potential a twisting of a warped gas layer is naturally caused. It is this twisting combined with dissipation that drives the settling process (Tohline et al. 1982). It was noted by Christodoulou & Tohline (1986) that the direction of the twist relative to the direction of rotation of the disk depends on the type of the quadrupole distortion of the potential. When the orbital period increases monotonically with radius and the precession frequency of particles decreases monotonically with radius, which is normally the case in the outer parts of galaxies, then the twist is prograde with respect to the galaxy’s rotation if the potential is oblate. On the other hand, the twist is retrograde if the potential is prolate. If the direction of the twist can be derived unambiguously from the observations one can get the shape of the underlying potential well. However, from HI observations one can generally not determine the near edge
of a galaxy and without knowing the near edge one cannot decide whether the twist is prograde or retrograde with respect to the rotation. This is called the orientation ambiguity by CTS. But optical photographs often show dust lanes in the equatorial plane of the galaxies, and from these optical images one can generally determine the near edge of a galaxy. The variation of the twist and the warp was modeled by CTS generally by simply increasing the twist and the warp angles linearly with radius. But recently, Barker et al. (1993) derived an analytic formula that describes the warp and twist of a disk as it settles in a spherical potential with a quadrupole distortion. In this chapter we will use galaxy models that they constructed from this analytic formula.

By putting the above described features together we can construct models of the HI disk. The distribution of neutral hydrogen is described by the function

\[ n(R) = n_0 \exp(-\Lambda R/R_0 - (z - z_0)/h) \]

where \( n_0 \) is the number density of neutral hydrogen at the center and \( \Lambda \) is the scale length of the exponential disk. We chose \( \Lambda = 1 \) such that the density at \( R_0 \) has decreased to \( 1/e \). Outside the radius \( R_0 \) the density is set to zero. The dependence of the density in the \( z \) direction is modeled with an exponential decrease with a scale height \( h \) from a local reference \( z_0 \), which in turn is a function of the radius and depends on the warp and the twist of the galaxy.

The HI gas moves on circular orbits and the velocity is modeled as

\[ V(R) = \begin{cases} (R/R_c)V_c, & R < R_c \\ V_c, & R_c \leq R \leq R_0 \end{cases} \]

where \( R_c \) is the core radius and \( V_c \) is the maximum rotational velocity. Typical values of \( V_c, R_0, \) and \( R_c \) in spiral galaxies are 250 km s\(^{-1}\), 30 kpc, and 5
kpc, respectively. Here we have not bothered to tune these parameters separately for each modeled system because we are concerned only with qualitative comparisons with observed velocity and surface density maps.

Following the traditional tilted-ring prescription of a warped galaxy disk (Rogstad et al. 1974; CTS) our model galaxy is constructed by using \( N \) rings whereby the \( n \)th ring has a warp \( w_n \) and the warp node is twisted with respect to some reference axis in the \( x-y \) plane by an angle \( t_n \). This reference axis is rotated from the \( x \) axis by an angle \( t_0 \), and the \( x-y \) plane is tilted around the \( x \) axis by an angle \( i_0 \) in the direction of the \( y \) axis. These angles are related to the Euler angles in our imaging program by \( \Theta = -i_0 \) and \( \Phi = -t_0 \). In the following we will simply set the third Euler angle \( \Psi \) to zero. Here we adopt warp and twist angle functions \( w_n(R) \) and \( t_n(R) \) based on the analytic formula by Barker et al. (1993). Put in formula we then have only two independent parameters that describe each model: the sign of the twist, and the amplitude of the warp at the outermost radius \( w_{\text{max}} \). Here, all our model galaxies have oblate potentials (i.e., the twist is prograde) and \( w_{\text{max}} \) varies over the range \( 4.8^\circ \leq w_{\text{max}} \leq 21.3^\circ \).

Once we have constructed the model galaxies we can use the imaging code to view them from different directions. We have produced a video of a "fly-by" around these model galaxies by smoothly changing the values of the viewing angles \( i_0 \) and \( t_0 \) and only occasionally changing \( w_{\text{max}} \). The calculated column density and velocity profiles are derived directly from the density and velocity distribution of matter, and the radiative transfer equation has not been solved. We therefore assumed that the radiation is optically thin and
that the observations were ideal (with infinite resolution and no noise). This is justified here as we are only making qualitative comparisons with the observations and do not attempt to match them in every detail. We have calculated the column density and radial velocity maps resulting from over 300 separate viewing directions and Howard Cohl made 3D image renderings of the model galaxies as seen from the respective directions using the graphic package IDL.

The angles used to define the view of a 3D object in IDL (denoted as $a_x$ and $a_z$) are related to our Euler angles by $a_x = -\Theta - \pi/2$ and $a_z = -\Phi$.

It has turned out that the appearance of a galaxy, i.e. the observed column density and radial velocity profiles, can change dramatically by simply varying the viewing direction. While the value of the warp angle is of course important particularly if one tries to match observations in detail, it is the viewing direction which most critically determines the qualitative appearance of a galaxy. We illustrate this by displaying in Figs. 3 to 8 the images that resemble six different observed galaxies. These images are taken directly from the video. We have not tried to make a best fit to the individual galaxies (this was done for example by CTS for a large number of galaxies). In the figure captions we list references to HI observations of the individual galaxies against which our modeled maps should be compared. For two galaxies (NGC 5033 and NGC 5055) we show the observed column density and radial velocity profiles from Bosma (1981). Figures 9 and 10 show the radial velocity and column density maps for NGC 5033, while Figs. 11 and 12 show the corresponding maps for NGC 5055. The main parameters of the model galaxies are listed in
Table 2. As we said above, all these model galaxies have the same structure and only the value of the warp angle changes slightly.

As Christodoulou & Tohline (1986) showed, the direction of rotation of a galaxy with respect to the direction of the twist determines the shape of the underlying potential. However, to determine this unambiguously one needs to know which is the near and which is the far side of a galaxy. Radio observations of neutral hydrogen cannot in general distinguish between the near and the far side of a galaxy because the radiation is optically thin. Many times, however, this can be determined from optical images which often show dust lanes. It is therefore the combination of optical and radio observations which determine the underlying potential. For illustration purposes, we have constructed disks with only prograde twists, that is, all our models have had oblate potentials. CTS determined the "correct" shape of the underlying potential for a number of galaxies (see Table 2). As can be seen by comparing our images to the actual observations, the images of all galaxies that have oblate potentials as determined by CTS match the observations. But the galaxy NGC 5033, which has a prolate potential according to CTS does not match with our images. If one aligns our theoretical velocity map for NGC 5033 with the observational one in Fig. 9 in such a way that the red shifts and the blue shifts match, then our theoretical 3D structure of the galaxy displays the near edge of the galaxy on the wrong side compared to the optical observations. This is expected as our models are based on oblate potentials and observations show the galaxy NGC 5033 to have a prolate potential. The same comparison as for NGC 5033 can be done with NGC 5055. CTS determined a prolate potential. Just as
Table 2. Model parameters for the different galaxies. The form of the underlying potential is from CTS

<table>
<thead>
<tr>
<th>Name</th>
<th>Fig. #</th>
<th>$i_0$</th>
<th>$t_0$</th>
<th>$w_{\text{max}}$</th>
<th>Potential (CTS)</th>
<th>Potential (Here)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NGC 2841</td>
<td>3</td>
<td>284°</td>
<td>240°</td>
<td>4.5</td>
<td>oblate</td>
<td>oblate</td>
</tr>
<tr>
<td>NGC 7331</td>
<td>4</td>
<td>255°</td>
<td>240°</td>
<td>4.5</td>
<td>oblate</td>
<td>oblate</td>
</tr>
<tr>
<td>M83</td>
<td>5</td>
<td>174°</td>
<td>277°</td>
<td>8.3</td>
<td>oblate</td>
<td>oblate</td>
</tr>
<tr>
<td>NGC 5033</td>
<td>6</td>
<td>125°</td>
<td>289°</td>
<td>12.0</td>
<td>prolate</td>
<td>oblate</td>
</tr>
<tr>
<td>NGC 5055</td>
<td>7</td>
<td>125°</td>
<td>435°</td>
<td>12.0</td>
<td>prolate</td>
<td>oblate</td>
</tr>
<tr>
<td>NGC 300</td>
<td>8</td>
<td>13°</td>
<td>469°</td>
<td>21.3</td>
<td>oblate</td>
<td>oblate</td>
</tr>
</tbody>
</table>

above one can align the theoretical velocity profiles with the observed ones in Fig. 11 and compare the theoretical 3D structure and the optical observations. Like for NGC 5033 we expect that these do not match, i.e. that the near side of the theoretical image corresponds to the far side of the optical image. This, however, is not the case. A comparison shows that the 3D structure as calculated by us from an oblate potential and the optical disk as observed do align. Therefore, NGC 5055 must have an oblate potential and not a prolate potential as stated by CTS. It should be pointed out that our images were extracted from a movie in which we continuously fly by our model galaxies. Therefore all our images are consistent in themselves, which leaves little room for an error.

In the following we want to summarize this study. The aim of this work has been to study qualitatively the appearance of a model galaxy when it is viewed from different angles. We varied the physical structure of our model galaxy by using different maximum warp angles $w_{\text{max}}$. However, it has
Figure 3. HI distribution in our model galaxy resembling NGC 2841. The figure shows the 3D structure (left), the density profile (middle) and the velocity profile (right) as calculated from our model galaxy. In the velocity profile, white corresponds to blue shift and black corresponds to red shift. For density and velocity profiles from observations of this galaxy see Bosma (1981)

turned out that the qualitative appearance of the column density and radial velocity profiles of a galaxy (which is what radio astronomers obtain from measurements of the 21 cm line) mainly depend on the viewing direction. The actual strength of the warp (as expressed by $w_{\text{max}}$) is only important when a detailed quantitative fit to an observed galaxy is made.
Figure 4. Same as Fig. 3 but for the galaxy NGC 7331. For density and velocity profiles from observations of this galaxy see Bosma (1981)

Figure 5. Same as Fig. 3 but for the galaxy M83. For density and velocity profiles from observations of this galaxy see Rogstad et al. (1974)
Figure 6. Same as Fig. 3 but for the galaxy NGC 5033. For density and velocity profiles from observations of this galaxy see Bosma (1981).

Figure 7. Same as Fig. 3 but for the galaxy NGC 5055. For density and velocity profiles from observations of this galaxy see Bosma (1981).
Figure 8. Same as Fig. 3 but for the galaxy NGC 300. For density and velocity profiles from observations of this galaxy see Rogstad et al. (1979)
Figure 9. Observational radial velocity map of NGC 5033 from Bosma (1981). The near edge is on the right (west)
Figure 10. Observational column density map of NGC 5033 from Bosma (1981)
Figure 11. Observational radial velocity map of NGC 5055 from Bosma (1981). The near edge is on the bottom (south)
Figure 12. Observational column density map of NGC 5055 from Bosma (1981)
CHAPTER 5

DUST IN OUTER REGIONS OF GALAXIES

5.1 Introduction

Some galaxies show strong radial asymmetries, that is there appears to be more visible matter on one side of the galaxy than on the other. Sometimes these asymmetries could have resulted from a perturbation by a nearby companion. But in other cases, no companion is seen, and the asymmetry seems to be persistent.

Asymmetries of the neutral hydrogen distribution have been investigated by Baldwin et al. (1980). They proposed that gas moves on elliptical orbits instead of circular orbits. Normally, elliptical orbits will be circularized due to the effects of shear. But at the radii of the outer HI regions the Hubble time is only a few times the rotation period, and the equilibrium may not yet have been reached. However, it is difficult to explain asymmetries of the optical disk with this mechanism as its size is much smaller and its rotation period therefore shorter than that of the outer HI disk.

In this chapter, we propose the following new mechanism for the apparent asymmetry in some optical disks. As has been discussed earlier in Chapter 4, HI observations have revealed strong warps in the outer regions of many disk galaxies. If a galaxy has in addition a high inclination to our line of sight, then the outer HI disk can cross over the optical disk, as seen by the observer. In the case that dust is present in these outer HI regions, the galaxy
may appear asymmetric. However, one would normally not expect any dust at these large distances, as no stars seem to be present there which could provide elements heavier than hydrogen and helium to form the dust. This hypothesis of presence of dust beyond the Holmberg radius, which is approximately the observed outer radius of the optical disk, must fulfill the following constraints. First, the model of the HI distribution must match the radio observations of neutral hydrogen. Second, the HI distribution must cross over the optical disk, and the asymmetry must appear where this cross-over occurs. Furthermore, the asymmetry must be strongest at short wavelengths and weakest at long wavelengths because of the wavelength dependence of the scattering cross section of dust.

We proceed as follows. First we describe the general features of the model galaxy, which is based on the tilted-ring model by Christodoulou et al. (1993, hereafter CTS). In this section we also discuss the radiative transfer. In Sect. 5.3 we discuss the galaxy NGC 5055. A radial asymmetry of this galaxy can be seen for example in its photograph by Sandage (1961; see also CTS).

5.2 Galaxy model and radiative transfer

To construct the HI disk of the galaxy we use the tilted-ring model of CTS. In this model the gas rotates on $N$ circular rings around the galaxy within the radius $R_0$ of the outer edge of the galaxy. The innermost regions form a flat disk of radius $R_{cd}$. The inclination of this central disk is given by the angle $i_0$ and is derived from optical observations. Outside $R_{cd}$ the orbits of gas become twisted and warped, and the position angle of the warp of the first ring outside $R_{cd}$ is $t_0$. To constrain the model, the twist and warp of the individual rings
are not determined independently but instead increase linearly outward. The increase in the twist and warp per ring is $\Delta t$ and $\Delta w$, respectively. The rotation of the gas moving on the rings is that of a rigid body within a radius $R_c$ and outside the velocity $V_c$ is constant as prescribed by Eq. (4.5). The maximum warp of the outermost ring is $w_{\text{max}}$ and the twist of this ring with respect to the first warped ring is $t_{\text{max}}$. If in addition, it is clear from optical observations which side the near side of the galaxy is and if the direction of the twist with respect to the rotation of the galaxy can be determined unambiguously, then the shape of the underlying potential, i.e. whether it is prolate or oblate, is known (Christodoulou & Tohline 1986). This in turn determines the side of the galaxy where the cross-over of the optical disk by the HI disk occurs.

In order to calculate the absorption due to dust in the outer regions of a galaxy we need to have the distribution of hydrogen, dust and stars. It is generally observed that galaxies have an exponential decrease of gas, dust and stars with radius in their outer parts. (The structure of galaxies at the center is more complicated, but this is not important here, as we are only concerned with the properties of matter far outside the center of galaxies.) Here we only want to reproduce qualitative features. In an analogous fashion to Eq. (4.4), we model the number densities $n_g$, $n_d$, $n_s$ of gas, dust and stars as follows.

$$
\begin{align*}
n_g &= n_g(0) \exp \left[ - \left( \frac{z - z_0}{h_g} \right) - \Lambda_g R / R_0 \right] \\
n_d &= n_d(0) \exp \left[ - \left( \frac{z - z_0}{h_d} \right) - \Lambda_d R / R_0 \right] \\
n_s &= n_s(0) \exp \left[ - \left( \frac{z - z_0}{h_s} \right) - \Lambda_s R / R_0 \right]
\end{align*}
$$

(5.1) (5.2) (5.3)

where $z_0$ is the distance of the surface of maximum density from the galactic plane, and $h_g$, $h_d$, $h_s$ are the scale heights for the different components (for
simplicity we assume that the scale heights are independent of the radius). The decrease of the densities with radius is determined by the parameters \( \Lambda_g, \Lambda_d, \Lambda_s \). Each component also has a maximum radius beyond which the density is set to zero. For hydrogen this is \( R_0 \), while for stars it is the Holmberg radius \( R_H \), which can be much smaller than \( R_0 \). The distribution of the dust is less certain as it depends on its uncertain formation. According to our hypothesis there is a component of the dust which has the same distribution as HI. Therefore \( \Lambda_g = \Lambda_d \) and its cutoff is \( R_0 \). Likewise, we assume \( h_g = h_d \).

For the stars we set \( h_s = h_g \). In the inner regions of galaxies \( h_s > h_g \). However, we are interested here primarily in the outer structure of galaxies. There the gas layer can become thick (see e.g. Sancisi & Allen 1979 for NGC 891 and Burton 1992 for our Galaxy), but we will not consider the radial dependence of the scale height here.

The image of galaxies at some wavelength \( \lambda \) when taking into account absorption due to dust is calculated by solving the radiative transfer equation in the form

\[
\frac{dI_\lambda}{ds} = -\chi_\lambda I_\lambda + \eta_\lambda
\]

where \( I \) is the specific intensity, \( s \) is the length along the line of sight, \( \chi \) is the absorption coefficient due to dust and \( \eta \) is the emissivity, which models the emission due to stars. Formally, the emissivity depends on the wavelength, but as we are only interested in radial asymmetries, i.e. intensity differences at the same wavelength and radius, we can regard \( \eta \) to be independent of \( \lambda \). At wavelengths larger than about 3000 Å one finds from measurements of the extinction \( A_\lambda \) that \( A_\lambda \propto \lambda^{-1} \). In the B band (\( \lambda = 4350 \) Å) the absorption is
therefore almost twice as strong as in the I band ($\lambda = 8250$ Å). By comparing the column density $N_\text{g}$ of atomic hydrogen and the values of interstellar reddening $E_{\text{B-V}}$ of individual stars at distances of 0.1 to 1 kpc one is led to the relation (Scheffler & Elsässer 1987)

$$N_\text{H}/E_{\text{B-V}} = 5.8 \times 10^{21} \text{ H atoms cm}^{-2} \text{ mag}^{-1}. \quad (5.5)$$

With a mean reddening of $E_{\text{B-V}} = 0.6$ mag kpc$^{-1}$ one obtains a number density of neutral hydrogen of $n_\text{g} = 1.1$ cm$^{-3}$. Also from the measured mean ratio of interstellar extinction to interstellar reddening one has $R = A_\text{V}/E_{\text{B-V}} \approx 3$. As $A_\lambda = 1.086\tau_\lambda$ where $\tau_\lambda$ is the optical depth one can therefore write

$$\chi_\lambda = \sigma_\text{V} \frac{\lambda_\text{V} n_\text{d}}{\lambda} n_\text{g}. \quad (5.6)$$

where the wavelength in the V band is $\lambda_\text{V} = 5550$ Å. The absorption cross section $\sigma_\text{V}$ depends on the size and the composition of the dust grains, and both are not known very accurately. The dust-to-gas ratio is also not known very well. A typical value is $n_\text{d}/n_\text{g} = 4 \times 10^{-13}$ (Lang 1986) and one obtains $\sigma_\text{V} = 1.2 \times 10^{-9}$ Å. This also agrees with Jura (1980), who gives an optical depth in the B band of $\tau_\text{B} = 6.4 \times 10^{-22} N_\text{H}$. From this one obtains a cross section of $\sigma_\text{B} = 1.6 \times 10^{-9}$ for $n_\text{d}/n_\text{g} = 4 \times 10^{-13}$. A similar value is derived by Lang (1986) from a simple estimate. He uses a size of dust grains of $a = 0.2 \times 10^{-4}$ cm and obtains a cross section $\sigma_\text{V} \approx \pi a^2 \approx 10^{-9}$ cm$^2$. In the following we will adopt values of $\sigma_\text{V} = 1.2 \times 10^{-9}$ cm$^2$ and $n_\text{d}/n_\text{g} = 4 \times 10^{-13}$. However, one should keep in mind that these values are still very uncertain. Furthermore, it is assumed that the composition of interstellar dust grains is the same everywhere, which may not be true.
The energy level of the electron of neutral hydrogen in the ground state is split depending on whether the spin of the electron has the same or the opposite orientation with respect to the spin of the nucleus. The transition between these two states gives rise to the 21 cm radiation of neutral hydrogen. To calculate the HI distribution from the 21 cm line radiation (see Mihalas & Binney 1981 for a more detailed discussion) we use the radiative transfer equation in the form

$$\frac{dI_\nu}{ds} = \chi_\nu (B_\nu(T_s) - I_\nu)$$ \hspace{1cm} (5.7)

where the Planck function $B$ (in the Rayleigh-Jeans limit) is used as the source function and $T_s$ is the spin temperature, which determines the distribution of the electron in the two spin states. It is approximately the same as the kinetic temperature of neutral hydrogen. The absorption coefficient is given as

$$\chi_\nu = (3hc^2/2\pi^2\lambda^3)n_H\Phi_\nu$$ \hspace{1cm} (5.8)

where $\Phi$ is the absorption profile. This profile is dominated by Doppler broadening and we can write

$$\chi_\nu = \frac{1}{\pi^{1/2}\Delta\nu_D} \exp \left[ - \left( \frac{\nu_0 - \nu - \nu_D v_r/c}{\Delta\nu_D} \right)^2 \right]$$ \hspace{1cm} (5.9)

with the Doppler width $\Delta\nu_D = \nu_D v_D/c$ and $v_r$ is the velocity along the line of sight. The characteristic velocity of the hydrogen atoms $v_D \approx 5$ km s$^{-1}$. We use Eqs. (5.7) to (5.9) to calculate the intensity of the 21 cm line of neutral hydrogen. However, we interpret the calculated image by assuming that the matter is optically thin and that the Planck function is constant along the line of sight. These assumptions are generally made by radio astronomers, and they
are justified here. Then the brightness temperature (see Eq. 4.2) becomes

\[ T_B \approx \tau_\nu T_s \]  

(5.10)

where \( \tau_\nu \) is the optical depth along the line of sight. The integrated column density along a line of sight is then given by integral (4.1) and is independent of the only poorly known spin temperature. When we calculate the column density map and radial velocity map of our model and compare them with observations we take into account the finite filter width for the frequency, the noise within each frequency interval and the finite spatial resolution of the observations.

5.3 NGC 5055

The HI distribution of this galaxy has been measured and modeled by Bosma (1981) and modeled separately by CTS, who could reproduce the main observed features. They determined that this galaxy has a prolate potential. We begin by adopting the best fit warped disk structure found by CTS. However, as we described in Chapter 4, the underlying potential of this galaxy actually must be oblate. CTS used an inclination of \( i_0 = 55^\circ \) and \( t_0 = 260^\circ \). Because we found that this galaxy has an oblate potential in order to implement the CTS "best fit" structure we must perform the transformations \( i_0 \rightarrow 180^\circ - i_0 = 125^\circ \) and \( t_0 \rightarrow 180^\circ - t_0 = -80^\circ \). But we change \( t_0 \) somewhat and use \( t_0 = -20^\circ \). CTS had a radius \( R_0 = 17'.6 \) (here we adopt a distance scale such that this value of \( R_0 \) corresponds to 30 kpc). However, we further increase \( R_0 \) by 40% and use a linearly varying warp and twist such that \( w_{\text{max}} = 38^\circ \) and \( t_{\text{max}} = -164^\circ \). The factor \( \Lambda_g \) is such that the HI density
Has decreased by a factor 10 at a radius of 30 kpc, just as in the model of CTS. The increase of \( R_0 \) is justified as the density of neutral hydrogen is too small to be detected at the sensitivity of the observations of Bosma (1981). We demonstrate this in Fig. 13a, in which we have calculated the surface density of HI of our model of the galaxy when taking into account the sensitivity of the radio observations while Fig. 13b shows the computed velocity profile of the galaxy. We calculated the specific intensities at different wavelengths along each ray, folded over the filter width of 27.2 km \( \text{s}^{-1} \) and the beam width of 49\(^\prime\) \( \times \) 72\(^\prime\) and then obtained the column densities by only summing over the intensities that were above the noise of 0.5 K. In these models we used a central hydrogen density of \( n_g(0) = 0.125 \text{ cm}^{-3} \) and a scale height of \( h_g = 0.75 \text{ kpc} \). When one compares these two figures to the observations by Bosma (1981; see Figs. 11 and 12) one sees, that our model can reproduce the main observed features. These include the closed velocity contours and the northern "bridge" of HI. In the density profile we use the same values for the column density as Bosma (1981). One can therefore see, that the extent of the modeled HI disk on the sky (after taking into account the limitations of the observations) is about the same as the observed one even though we have increased the radius of the disk by 40 \%. One should also take into account that our model uses a perfectly smooth distribution of HI. In reality the HI distribution is much noisier. This will rather decrease the amount of HI observed in the outer regions when one takes the noise of the observations into account. What our model does not reproduce entirely is the distribution of HI in the southern part of the galaxy. There the distribution of neutral hydrogen becomes much more irregular as
compared to the northern region. This suggests that the gas in the southern region is not in the quasi steady state that is characteristic for gas settling in an oblate potential.

The three-dimensional structure of our model galaxy is displayed in Fig. 14. For reasons of clarity we only display the HI disk out to the radius where the HI disk starts to cross over the (inner) optical disk. The HI disk first obscures the inner regions in the SW edge of the galaxy. From this it becomes clear how and where a radial asymmetry would occur if dust existed in the outer regions of the HI disk. We illustrate this further by calculating the image of the optical disk when there is dust in the HI disk. For this we solve Eq. (5.4) for wavelengths corresponding to the center of the B and the I band. The images are displayed in Figs. 15a and b. These images shall only reflect the general qualitative behavior of the light distribution of the galaxy. No detailed quantitative modeling has been attempted. As the effect of absorption due to dust in the outer HI regions is weak we increase the amount of dust by a factor 6 in order to illustrate more clearly the effects of dust. As one can see, parts of the optical disk are partially obscured by the dust in the extended HI disk, which causes an apparent asymmetry of the optical disk. This asymmetry reduces towards longer wavelengths because of the decreasing absorption coefficient.

The optical images of our model should be compared to the CCD observations of the galaxy NGC 5055. In Figs. 16 and 17 we show CCD images of the galaxy NGC 5055 in the B and the I band, respectively (these images were provided to us by D. Elmegreen, private communication). When one tries
Figure 13. Calculated HI map of NGC 5055. (a) Column density and (b) velocity profile of our model galaxy when the observational constraints are taken into account. In (a) the lowest contour (dashed) corresponds to a column density of $1.3 \times 10^{20}$ atoms cm$^{-2}$. All other contours (solid) are multiples of $2.6 \times 10^{20}$ atoms cm$^{-2}$. In (b) the contours are in intervals of 20 km s$^{-1}$. The blue shifted side is on the right. The zero velocity contour is dashed.
to determine quantitatively the amount of extinction caused by the dust, one needs to perform azimuthal scans of the galaxy, i.e. one needs to measure the light originating from the galaxy at a constant radius from its center in the plane of the galaxy. For this one has to know the position angle of this galaxy in the sky. However, this is determined by the contour lines of the galaxy which in turn are influenced by the radial asymmetry. CTS gave a position angle of $\gamma = 99^\circ$. On the other hand, Elmegreen & Elmegreen (1984) determined $\gamma = 105^\circ$. Though this may not appear to be a large difference, it influences greatly the measured amount of extinction. To determine quantitatively the amount of dust in the outer regions of the galaxy we compare the photon count at a fixed radius between two different quadrants, one where we expect the cross-over to obscure the optical disk (the SW quadrant) and one where the disk should not be obscured (the SE quadrant). In order to reduce the error we use a long azimuthal strip per quadrant instead of only one point.
Figure 15. Optical image of our model galaxy. (a) at 4350 Å, (b) at 8250 Å. The upper images are for the case of dust in the HI disk while the lower images are for the case that dust only exists in the optical disk.
Table 3. Dust in the outer regions of NGC 5055. For different position angles $\gamma$ we list the measured extinction $A_{\lambda}$ in the B band ($\lambda = 4350$ Å) and I band ($\lambda = 8250$ Å), the resulting column density $N_d$ of dust, and the dust-to-gas ratio assuming a column density of gas of $N_g = 10^{20}$ cm$^{-2}$.

<table>
<thead>
<tr>
<th>$\gamma$ [°]</th>
<th>$\lambda$ [Å]</th>
<th>$A_{\lambda}$</th>
<th>$N_d$ [cm$^{-2}$]</th>
<th>$N_d/N_g$</th>
</tr>
</thead>
<tbody>
<tr>
<td>99</td>
<td>4350</td>
<td>0.19 ± 0.12</td>
<td>$(1.1 \pm 0.7) \times 10^8$</td>
<td>$(1.1 \pm 0.7) \times 10^{-12}$</td>
</tr>
<tr>
<td></td>
<td>8250</td>
<td>0.11 ± 0.05</td>
<td>$(1.3 \pm 0.5) \times 10^8$</td>
<td>$(1.3 \pm 0.5) \times 10^{-12}$</td>
</tr>
<tr>
<td>102</td>
<td>4350</td>
<td>0.086 ± 0.037</td>
<td>$(5.2 \pm 2.3) \times 10^7$</td>
<td>$(5.2 \pm 2.3) \times 10^{-13}$</td>
</tr>
<tr>
<td></td>
<td>8250</td>
<td>0.012 ± 0.025</td>
<td>$(1.4 \pm 3.0) \times 10^7$</td>
<td>$(1.4 \pm 3.0) \times 10^{-13}$</td>
</tr>
<tr>
<td>105</td>
<td>4350</td>
<td>−0.01 ± 0.005</td>
<td>$(6.5 \pm 3.1) \times 10^6$</td>
<td>$(6.5 \pm 3.1) \times 10^{-14}$</td>
</tr>
<tr>
<td></td>
<td>8250</td>
<td>−0.1 ± 0.025</td>
<td>$(1.2 \pm 0.3) \times 10^8$</td>
<td>$(1.2 \pm 0.3) \times 10^{-12}$</td>
</tr>
</tbody>
</table>

per quadrant. These strips are shown in Figs. 16 and 17 for a position angle of $\gamma = 102^\circ$. Table 5.3 lists the results. The error estimate is based on the mean deviation of the photon count along each radial strip. While for $\gamma = 99^\circ$ we get large absorption in the SW quadrant, which is what we expect from our hypothesis, our measurements result in absorption in the SE quadrant (indicated by the negative extinction) for $\gamma = 105^\circ$. This is surprising as from optical photographs (Sandage 1961; CTS) the radial asymmetry appears to be fairly clear with less light coming from the SW quadrant than the SE quadrant. One problem of such radial scans are certainly the small scale structures of the galaxy which produce a large error. Therefore we cannot obtain a quantitative estimate of the amount of dust present in the outer HI regions.
Figure 16. Contour plot of a CCD image of NGC 5055 in the B band. North is up and West is on the right. The two stripes in the SE and SW mark the positions where the photon count for a quantitative determination of the dust absorption has been measured assuming a position angle of 102°. The CCD image was provided by Elmegreen (private communication)
Figure 17. Same as the Fig. 16 but this image of NGC 5055 is in the I band
5.4 Summary

In this chapter we showed how the three-dimensional structure of the HI disk could explain the observed asymmetries in optical galaxies if the outer regions of the HI disk contained dust. The advantage of this hypothesis is that no special assumption on the distribution of stars needs to be made. In the following, we want to summarize the basic features of this hypothesis.

1. The tilted-ring model has very few free parameters. There are only four parameters which determine the kinematics of the galaxy, and these are normally fairly well constrained by the observation of the velocity and density profile. There is an intrinsic ambiguity in HI observations. As this radiation is optically thin one cannot determine from these observations alone which the near and which the far side of a galaxy is. In our case this means that one cannot determine from the HI observations alone on which side we expect the optical disk to be obscured by the dust in the HI disk. However, by taking into account optical photographs of a galaxy one can usually determine the near side of a galaxy (except when the galaxy is viewed nearly face-on). This then determines where the HI disk crosses over. This theory can therefore predict where the radial asymmetry occurs.

2. Another important free parameter is the e-folding radius of the HI density. It can be fairly well determined from the HI density profile of the galaxy. For NGC 5055 we have to assume that the HI gas extends to larger radii than previously has been seen observationally. However, we have shown that the observed HI radius may merely reflect the detection limit of the radio observations of this galaxy and not the actual physical size of the HI disk.
3. The model predicts that the strength of the asymmetry is color-dependent. However, while a radial asymmetry is apparent in the photograph of NGC 5055 by Sandage (1961; CTS) we cannot determine quantitatively the extinction that is caused by the dust because of the large uncertainties associated with the required measurements at optical wavelengths.

4. A key assumption of this hypothesis is the presence of dust in the outer regions of the HI disk. There are currently several theories on where dust can form. Such formation regions are mainly the envelopes of protostars, atmospheres of cool giant stars and the general interstellar gas (Scheffler & Elsässer 1987). Even in the latter case stars are needed to produce the metals (particularly carbon) from which the dust is formed. However, no stars are observed in the outer regions, but the existence of dust at such radii is still possible. The metals from which dust forms could have been produced by stars formed during the initial star formation in the galaxy. As one knows from the halo component of our Galaxy this star formation was not yet confined to the disk of the galaxy. Within a short time the more massive stars in the outer halo ended their lives either through supernovae or after phases of intense stellar winds as white dwarfs. In either case they would have provided metals to the surrounding medium. Only the less massive and therefore unobservable stars were left behind. The metals then would settle in the same warped disk as neutral hydrogen atoms. There the metals could have condensed into dust.
CHAPTER 6

POLARIZED RADIATIVE TRANSFER

6.1 Introduction

Cataclysmic variables (CVs) are binary systems in which a white dwarf accretes matter from a close companion of low mass (0.1 to 1 $M_\odot$) which loses mass through Roche lobe overflow (Frank et al. 1992). Because of conservation of angular momentum, this matter (which is collisionally ionized) forms an accretion disk around the white dwarf and slowly spirals towards the white dwarf because of the influence of viscosity. However, when the white dwarf possesses a strong magnetic field, this field will disrupt the accretion disk and the plasma will flow along field lines onto the magnetic pole of the white dwarf (see review by Chanmugam 1992). Above the accretion region on the white dwarf, a shock will form when the plasma has to slow down from supersonic free fall velocities to essentially zero velocity on the surface of the white dwarf.

From observations, two classes of magnetic CVs have been distinguished (e.g. Chanmugam 1992). On the one hand there are the synchronous systems in which the spin period of the white dwarf is about equal to the orbital period, i.e. $P_{\text{spin}}/P_{\text{orbit}} \approx 1$. These systems are also known as AM Her stars or polars (Cropper 1990). They display strong (about 10 %) phase-dependent linear and circular polarization in the optical and infra-red. From this it follows that the white dwarf has a strong magnetic field of about 10 to 70 MG. In addition, all of these stars have a short orbital period (< 4 hours). It is
this short period combined with the strong magnetic field, which causes the synchronous rotation of the white dwarf. As a consequence, all emission from the x-ray to the infrared varies with $P_{\text{spin}}$.

A second class are the so-called asynchronous magnetic CVs in which the spin period is much shorter than the orbital period. This class of systems has previously been subdivided. Systems with $P_{\text{spin}}/P_{\text{orbit}} \sim 0.1$ are known as intermediate polars (IPs, e.g. Warner 1985). In these systems the hard x-ray emission is pulsed at $P_{\text{spin}}$. If the interpretation of this as optically thin bremsstrahlung is correct, the accretion occurs onto a large area (King & Shaviv 1984). Until recently no polarization of the emitted light was seen, and there was therefore no direct evidence of a magnetic field. This led people to believe that the magnetic field must be weak ($\ll 1$ MG). Besides the IPs the small class of DQ Her systems has been identified. These systems have similar properties as the IPs but $P_{\text{spin}}/P_{\text{orbit}} \ll 0.1$ and the x-ray emission is weak or completely absent.

It was long noticed that in general the orbital period of the AM Her objects is shorter than the period of the DQ Hers and IPs. This suggests an evolutionary sequence, in which the IPs evolve into the AM Her objects. However, for such a sequence to occur one needs some IPs with fairly strong magnetic fields ($B > 3$ MG; Chanmugam & Ray 1984). An evolution from IPs into AM Her stars was also suggested by King et al. (1985), who suggested that these two systems have magnetic fields of the same order of magnitude. This raises the question why no polarization is observed in IPs, and it has been suggested that this is due to dilution effects. Also if IPs evolve into AM
Her systems, then one would expect to find some systems which are in an evolutionary state in between IPs and AM Hers. The evolutionary scenario was supported by the discovery of BG CMi by Penning et al. (1986). This object has a spin period of 15.3 min and an orbital period of 3.23 hr and was therefore classified as an IP. In addition it showed weak (0.239 %) circular polarization in the I band (centered at 8250 Å), which provided the first direct evidence of a magnetic field in an IP. The calculations of Chanmugam et al. (1990) gave a field strength of \( \approx 4 \) MG.

Recently, the IP RE 0751+14 was discovered by Mason et al. (1992) in the EUV with the ROSAT satellite, which may also support the evolutionary scenario. This system showed a 13.9 min modulation, which they interpreted as the spin period. Rosen et al. (1993) obtained optical spectroscopy, photometry and polarimetry of this object. They noticed a radial velocity motion of the emission lines which supported a spin period of 13.9 min. In the B band they found a 14.5 min modulation, which can be interpreted as a beat period with the orbital motion \( P_{\text{orbit}} = 5.3 \) hr. They also found a strong red color dependence of the pulsation, which suggests a cyclotron origin. Furthermore, they noticed variable circular polarization. Both hint at a strong magnetic field. Finally, Pirola et al. (1993) made simultaneous UBVRI, linear and circular polarimetry, and they found variable linear and circular polarization. By fitting the Wickramasinghe & Meggitt (1985) cyclotron emission model to the maximum polarized flux they obtained a magnetic field of \( B = 8 - 18 \) MG with a temperature of \( kT = 10 - 20 \) keV. A third important parameter is the plasma parameter \( \Lambda = \omega_p^2 L/\omega_c c = 6 \times 10^7 (L/10^7 \text{ cm}) (N/10^{16} \text{ cm}^{-3})(B/10 \text{ MG}) \).
Here \( \omega_p = (4\pi N e^2/m)^{1/2} = 5.64 \times 10^{12} (N/10^{16} \text{ cm}^{-3})^{1/2} \text{ s}^{-1} \) is the plasma frequency, \( \omega_c = eB/mc = 1.76 \times 10^{14} (B/10 \text{ MG}) \text{ s}^{-1} \) is the cyclotron frequency, \( N \) is the electron number density and \( L \) is the characteristic length of the system. Pirola et al. (1993) found \( \Lambda \approx 10^6 - 10^8 \).

We have used our code to calculate polarized radiative transfer in a 3D emission region (the accretion shock) to polarized radiative transfer. In Appendix A we discuss all aspects of polarized radiative transfer as far as they are relevant here. There, we discuss both radiative transfer in the general Stokes formalism and radiative transfer in the special case of large Faraday rotation and pulsation. In Sect. 6.2 we discuss the numerical calculation of the required cyclotron absorption coefficients. This discussion is quite extensive. The reason for this is the lack of such a discussion in the literature for the opacities needed here (Väth & Chanmugam 1994). Then we calculate the spectra of objects with homogeneous magnetic fields and compare our results with those of Wu (1989). The assumption of homogeneous fields is realistic for an AM Her system, where matter only accretes onto a small fraction of the surface of the white dwarf. In the case of IPs, however, one cannot make this assumption and a more general calculation must be performed. We apply our program to the case of extended polar caps and we model the IP RE 0751+14.

### 6.2 Cyclotron opacities

#### 6.2.1 Introduction

Cyclotron radiation is produced by electrons as they are spiraling in a magnetic field (e.g. Bekefi 1966). This radiation is called synchrotron radiation when the electrons move at ultra-relativistic speeds. The collective emission of
electrons in a magnetized plasma depends on the electrons' velocity distribution. Here we will assume that the velocity distribution is Maxwellian including relativistic corrections. Then the cyclotron emission of a plasma is described by the temperature $T$, the angular frequency of emission $\omega$ parametrized by the harmonic number $s = \omega/\omega_c$, the angle $\theta$ between the magnetic field and the line of sight, and finally the plasma parameter $\Lambda$. If cyclotron emission is an important source of radiation then by Kirchhoff's law it is also a strong absorber of radiation. Therefore cyclotron absorption is important for the transfer of radiation through a strongly magnetized plasma.

Cyclotron emission can be detected from many astronomical systems. Examples are the thermal and non-thermal radio emission of the sun and stars (e.g., Dulk 1985). Cataclysmic variables containing accreting magnetic white dwarfs are also a source of strongly polarized optical and near-infrared radiation, whereby the magnetic field has a strength of $B \approx 10^7$ G (Chanmugam & Dulk 1981; Meggitt & Wickramasinghe 1982). X-ray cyclotron lines are seen in accreting neutron stars which are consistent with magnetic fields of $B \approx 10^{12}$ G (Trümper et al. 1978; Zheleznyakov 1984).

Extensive discussions of cyclotron opacities exist in plasma physics (e.g., Bekefi 1966; Tamor 1977, 1978; De Barbieri 1980; Bornatici et al. 1983) and in astrophysics (e.g., Ramaty 1969; Pavlov et al. 1980; Chanmugam & Dulk 1981; Meggitt & Wickramasinghe 1982). In these papers two different methods were used. In the single-particle method the radiation emitted by one electron in a magnetic field was calculated and from this the emissivity of a collection of thermal electrons was obtained. The cyclotron opacities were
then derived using Kirchhoff's law (e.g., Bekefi 1966). On the other hand one has the dielectric tensor method in which the opacities are obtained from the complex dielectric tensor (e.g., Tamor 1977, 1978). However, the opacities were mostly presented in the form of figures. As opacities vary over many orders of magnitude for different harmonic numbers it is difficult to compare opacities calculated by different authors, as differences of order unity easily remain undetectable in such figures. It is therefore important to compare numerical tables of opacities to decide on the accuracy of different numerical methods.

Polarized radiation can be described by the four Stokes parameters $I, Q, U, V$. The corresponding transfer equation for polarized radiation contains three independent opacities commonly denoted as $\kappa, \eta$ and $\upsilon$ plus the two Faraday mixing coefficients $f$ and $h$ (see also Appendix A). Alternatively, the propagation of an electromagnetic wave through a plasma can be described by two orthogonal waves referred to as the ordinary (+) and the extraordinary (−) wave. Under the assumption of large Faraday rotation and Faraday pulsation and homogeneous fields along the line of sight one obtains two uncoupled transfer equations for these two modes, and the transfer is described by only two opacities $\alpha_\pm$. As the condition of large Faraday rotation and pulsation and homogeneous fields is often fulfilled in laboratory and astrophysical plasmas it has been common in the past to use the transfer equations for the ordinary and extraordinary mode. Chanmugam et al. (1989) made thorough comparisons between different numerical methods and analytic approximations for the opacities $\alpha_\pm$ and also produced extensive tables of opacities.
In the following we extend the work done by Chanmugam et al. (1989) to the opacities $\kappa$, $q$, and $v$. That way we are able to solve polarized radiative transfer in cases where the magnetic field is inhomogeneous along the line of sight. This will be necessary in some of the systems investigated here. This discussion is also needed as Chanmugam et al. (1989) compared their total opacity $\kappa$ to those used by Meggitt & Wickramasinghe (1982), who solved the transfer equation in the Stokes formalism, and discrepancies as large as 20% appeared. The following detailed discussion should help to resolve questions about the numerical accuracy of these calculations. We use the single-particle method to calculate the opacities. We test our opacities by calculating from them the opacities in the ordinary and extraordinary mode $\alpha_\pm$ and comparing those to other people’s results. Furthermore, we derive formulae to obtain $\kappa$, $q$ and $v$ from $\alpha_+$. That way we can use analytic expressions for $\alpha_\pm$ in order to get analytic expressions for $\kappa$, $q$ and $v$. In particular, we use the analytic formulae by Robinson & Melrose (1984; Dulk 1985) and compare the so obtained analytic expressions of $\kappa$, $q$ and $v$ with the numerical results. Extensive tables of these cyclotron opacities for different $T$, $\theta$, and $s$ are given in Väth & Chanmugam (1994).

6.2.2 Cyclotron opacities

Polarized radiation can be described by the Stokes parameters $I$, $Q$, $U$, $V$ (Rybicki & Lightman 1979). The total intensity is given by $I$, while the degree of polarization is $(Q^2 + U^2 + V^2)^{1/2}/I$. The degree of linear polarization is $(Q^2 + U^2)^{1/2}/I$ and the degree of circular polarization is $V/I$. The position angle $\chi$ for linear polarization is given by $\tan 2\chi = U/Q$. When the magnetic
field \( \mathbf{B} \) is lying in the y-z plane with the z-axis in the direction of propagation (in which case the emissivity corresponding to the Stokes parameter \( U \) \( \epsilon_U = 0 \)) then the transfer equation can be written as (Pacholczyk 1977)

\[
\frac{d}{dl} \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \begin{pmatrix} \epsilon_I \\ \epsilon_Q \\ 0 \\ \epsilon_V \end{pmatrix} + \begin{pmatrix} -\kappa & -q & 0 & -v \\ -q & -\kappa & f & 0 \\ 0 & f & -\kappa & -h \\ -v & 0 & h & -\kappa \end{pmatrix} \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix}
\]

(6.1)

where \( l \) is the path length. Here the Faraday mixing coefficients are

\[
f = (\omega_p^2/\omega_c) \cos \theta/(\omega^2/\omega_c^2 - 1)
\]

(6.2)

\[
h = (\omega_p^2/\omega_c) \sin^2 \theta/2(\omega^3/\omega_c^3 - \omega/\omega_c).
\]

(6.3)

The quantities \( \epsilon_I \), \( \epsilon_Q \), \( \epsilon_V \) are the emissivities for the respective Stokes parameters. The solution to this equation in the case that all coefficients are constant along the line of sight is (see Appendix A, Eqs. A.31 to A.34)

\[
I = I_p + \{A_1 \cosh \lambda l + A_2 \sinh \lambda l + A_3 \cos \mu l + A_4 \sin \mu l\} e^{-\kappa l}
\]

(6.4)

\[
Q = I_Q + \left\{ -\frac{q + f a_1}{\lambda} [A_1 \sinh \lambda l + A_2 \cosh \lambda l] \\
+ \frac{-q + f a_3}{\mu} [A_3 \sin \mu l - A_4 \cos \mu l] \right\} e^{-\kappa l}
\]

(6.5)

\[
U = I_U + \left\{ -a_1 [A_1 \cosh \lambda l + A_2 \sinh \lambda l] \\
- a_3 [A_3 \cos \mu l + A_4 \sin \mu l] \right\} e^{-\kappa l}
\]

(6.6)

\[
V = V_V + \left\{ -\frac{v - ha_1}{\lambda} [A_1 \sinh \lambda l + A_2 \cosh \lambda l] \\
+ \frac{-v - ha_3}{\mu} [A_3 \sin \mu l - A_4 \cos \mu l] \right\} e^{-\kappa l}.
\]

(6.7)

All the constants contain the cyclotron absorption coefficients and the Faraday mixing coefficients. In addition, the constants \( A_i \) contain the incoming radiation and the constants \( X_P \) where \( X \) stands for any one of the Stokes parameters.
These latter constants in turn also depend on the emissivities $\epsilon_X$. We derive the above solution in Appendix A, where we also define all the constants.

Here we are interested in the case of cyclotron emission and absorption in a homogeneous collisionless plasma with a uniform magnetic field $\mathbf{B}$. The velocity distribution of electrons of number density $N$ is assumed to be Maxwellian including relativistic corrections. Then the emissivity for the Stokes parameter $I$ is given as (Meggitt & Wickramasinghe 1982)

$$
\epsilon_I = \frac{Ne^2\mu}{2eK_2(\mu)} \int_{-1}^{1} d\beta_\parallel \sum_{n=n_0}^{\infty} F_I \exp(-\mu\gamma)(\omega\gamma^4/n)
$$

(6.8)

where

$$
F_I = (\cot \theta - \beta_\parallel \csc \theta)^2 J_n^2(n\xi) + \beta_\perp^2 J_n^{1'}(n\xi).
$$

(6.9)

Here, $\mu = mc^2/kT$, $\beta = v/c$ is the dimensionless velocity, $c\beta_\parallel$ and $c\beta_\perp$ are the velocities of the electron parallel and perpendicular to the magnetic field, respectively, $\xi = \beta_\perp \sin \theta/(1 - \beta_\parallel \cos \theta)$, $\gamma = (1 - \beta^2)^{-1/2}$. The harmonic number is related to $\gamma$ through the equation $\gamma = n/s(1 - \beta_\parallel \cos \theta)$. Also $n_0$ is the minimum value of $n$ such that $\beta_\perp^2 > 0$. The emissivities $\epsilon_Q$, $\epsilon_V$ for the Stokes parameters $Q$ and $V$ are obtained from the above equation for $\epsilon_I$ by substituting $F_I$ with $F_Q$ and $F_V$, respectively. The functions $F_Q$ and $F_V$ are defined as

$$
F_Q = (\cot \theta - \beta_\parallel \csc \theta)^2 J_n^2(n\xi) - \beta_\perp^2 J_n^{1'}(n\xi)
$$

(6.10)

$$
F_V = -2(\cot \theta - \beta_\parallel \csc \theta)J_n(n\xi)\beta_\perp J_n^{1'}(n\xi)
$$

(6.11)

As before we assume that the axes are defined such that the $z$ axis points toward the observer and the uniform magnetic field lies in the $y$-$z$ plane making the angle $\theta$ with the $z$ axis. One obtains the opacities using Kirchhoff's law, and
one has \( \kappa = \epsilon_1 / B, \ q = \epsilon_Q / B, \ v = \epsilon_V / B \) where \( B = kT \omega^2 / 4\pi^3 c^2 \) is the Planck function in the Rayleigh-Jeans limit. We can use this limit because \( h\omega / kT \ll 1 \) here.

The propagation of electromagnetic waves in a magnetoactive plasma can be described by two normal modes (Ramaty 1969), which are the ordinary mode (+) and the extraordinary mode (−). The intensity in the respective modes is denoted by \( I_{\pm} \). Let us denote with \( E_x, E_\theta \) the transverse components of the electric field vector of the electromagnetic wave, where \( E_\theta \) lies in the plane of the magnetic field \( B \) and the direction of propagation while \( E_x \) is perpendicular to this plane. Then the polarization coefficients of the ordinary and extraordinary mode are given by \( a_{\pm} = -i(E_x / E_\theta)_{\pm} \) and one gets

\[
\frac{a_{\pm}(s, \theta)}{2s \cos \theta} = -\sin \theta \pm \left[ \sin^4 \theta + 4s^2 \cos^2 \theta \right]^{1/2}.
\] (6.12)

To fully describe the radiation one needs to know the phase difference between these two modes. However, in the case that this phase difference is completely randomized, which is referred to as large Faraday rotation and Faraday pulsation, the equations of transfer in a magnetoactive plasma simplify and can be written as (e.g. Ramaty 1969)

\[
dI_{\pm} / dl = \alpha_{\pm} (B / 2 - I_{\pm}).
\] (6.13)

where \( B \) is the Planck function. The formal solutions of these two equations in the case that all coefficients are constant along the line of sight is

\[
I_{\pm} = I_{\pm}^0 \exp(-\alpha_{\pm} l) + (B / 2) [1 - \exp(-\alpha_{\pm} l)].
\] (6.14)

The incoming radiation is denoted by \( I_{\pm}^0 \). The Stokes parameters are calculated
The opacities $\alpha_{\pm}$ for these two modes are (Chanmugam & Dulk 1981; the refractive index is set to one, which is a very good approximation)

$$
\alpha_{\pm} = \frac{1}{B c K_{2}(\mu)} \int_{-1}^{1} d\beta_{\parallel} \sum_{n=0}^{\infty} \frac{1}{1 + a_{\pm}^{2}} \left\{ -\beta_{\perp} J_{n}(\mu) + \left[ a_{\pm} (\cot \theta - \beta_{\parallel} \csc \theta) J_{n}(\mu) \right] \right\}^{2} \exp(-\mu \gamma)(\omega \gamma^{4}/n). \quad (6.19)
$$

In this equation $B$ is the Planck function.

Because much computing time is required to calculate the Bessel function $J_{n}$ of high order and the derivative $J_{n}'$, it is common to use an approximation given by Wild & Hill (1971; Meggitt & Wickramasinghe 1982) for these functions. The accuracy of this approximation was discussed by Chanmugam et al. (1989). They found that an error not larger than about 2% results in the opacities. This approximation is therefore sufficient for radiative transfer calculations. However, as a reference we calculate all opacities in this chapter using more precise numerical codes for the Bessel functions.

When one has calculated the opacities $\kappa$, $q$ and $v$ one can obtain the opacities $\alpha_{\pm}$ for the ordinary and extraordinary mode using the equation

$$
\alpha_{\pm} = \kappa - \frac{1 - a_{\pm}^{2}}{1 + a_{\pm}^{2}} q + \frac{2a_{\pm}}{1 + a_{\pm}^{2}} v. \quad (6.20)
$$
As far as we are aware this relation has not been shown before. There also exists the relation

\[ \kappa = (1/2)(\alpha_+ + \alpha_-) \]  

(6.21)

In the following we will prove these two relations. For this purpose we define

\[ A_n = (\cot \theta - \beta \parallel) J_n(n \xi) \]  
\[ B_n = \beta \perp J'_n(n \xi) \]

Now we insert the expressions of \( \kappa, q \) and \( v \) into Eq. (6.20), and as \( a \pm \) does not depend on the integration variable \( \beta \parallel \) and the summation index \( n \) we can write

\[ \alpha_\pm = \kappa - \frac{1 - a_\pm^2}{1 + a_\pm^2} q + \frac{2a_\pm}{1 + a_\pm^2} v \]

(6.22)

\[ = \frac{1}{B 2cK_2(\mu)} \int_{-1}^1 d\beta \parallel \sum_{n=1}^{\infty} \beta \parallel \sum_{n=1}^{\infty} \frac{1}{1 + a_\pm^2} \left\{ (A_n^2 + B_n^2) - \frac{1 - a_\pm^2}{1 + a_\pm^2} (A_n^2 - B_n^2) \right\} e^{-\kappa(\omega s \gamma^4/n)} \]

(6.23)

\[ = \frac{1}{B 2cK_2(\mu)} \int_{-1}^1 d\beta \parallel \sum_{n=1}^{\infty} \frac{1}{1 + a_\pm^2} \left\{ a_\pm^2 A_n^2 + B_n^2 - 2a_\pm A_n B_n \right\} e^{-\kappa(\omega s \gamma^4/n)} \]

(6.24)

The right hand side agrees with Eq. 6.19. Now we want to prove Eq. (6.21). By inserting the expressions of \( \alpha_\pm \) into Eq. (6.21) and using the same arguments as above we derive

\[ \kappa = (1/2)(\alpha_+ + \alpha_-) \]

(6.25)

\[ = \frac{1}{B 2cK_2(\mu)} \int_{-1}^1 d\beta \parallel \sum_{n=1}^{\infty} \frac{1}{1 + a_\pm^2} \left\{ a_\pm^2 A_n^2 - 2a_+ A_n B_n + B_n^2 \right\} e^{-\kappa(\omega s \gamma^4/n)} \]

(6.26)

\[ = \frac{1}{B 2cK_2(\mu)} \int_{-1}^1 d\beta \parallel \sum_{n=1}^{\infty} \left\{ A_n^2 + B_n^2 \right\} e^{-\kappa(\omega s \gamma^4/n)} \]

(6.27)

In going from Eq. (6.26) to (6.27) we used the relation \( a_+ a_- = -1 \) which can be shown easily from Eq. (6.12). While these are formal proofs, we also can
deduce Eqs. (6.20) and (6.21) intuitively by comparing the formal solutions to
the respective transfer equations. We present this derivation in Appendix A.
Equation (6.21) is well known and we included the formal proof only for the
sake of completeness.

While we can get \( \alpha_{\pm} \) from \( \kappa \), \( q \) and \( v \) we need one more equation
to do the reverse step. However, information (the phase relation between the
modes) is lost when one goes from the Stokes parameters to \( I_{\pm} \). Therefore no
such third relation exists. Nevertheless we can derive an approximately valid
relation. If the radiation is completely polarized one has \( I = \sqrt{Q^2 + V^2} \). In the
optically thin case with no Faraday rotation or pulsation we have \( I/B = \kappa l \),
\( Q/B = ql \), \( V/B = vl \) where \( l \) is the path length. It then follows that

\[
\kappa^2 \approx q^2 + v^2. \tag{6.28}
\]

However, it turns out that the radiation is never completely polarized (Rybicki
& Lightman 1979; only monochromatic waves are 100 % polarized, but because
of the uncertainty principle such waves cannot exist) and Eq. 6.28 is only valid
approximately. As the opacities are atomic quantities Eq. (6.28) remains valid
with the same accuracy even in the case of large Faraday rotation and if the
radiation is optically thick. Using Eq. (6.28) we can calculate \( q \) and \( v \) as
follows:

\[
q \approx \left\{ a \Delta - |b| \sqrt{\kappa^2 - \Delta^2} \right\} \tag{6.29}
\]

\[
v \approx \left\{ -b \Delta + \text{sgn}(b) |a| \sqrt{\kappa^2 - \Delta^2} \right\} \tag{6.30}
\]

where \( \Delta = (\alpha_- - \alpha_+)/2 \). We also defined the quantities
\( a = (1 - a^2_+)/ (1 + a^2_+) \)
and \( b = 2a_+/\left(1 + a_+^2\right) \). We will discuss the validity of the above equations further down.

6.2.3 Discussion

We start the discussion by first investigating the accuracy of our numerical calculations. For this purpose, we compare in Table 4 the opacities \( \alpha_\pm \) obtained using the dielectric tensor method of Tamor (1977, 1978) to \( \alpha_\pm \) as calculated by us from opacities in the Stokes formalism and using Eq. (6.20). This comparison is important as the methods used are qualitatively very different. While our calculation is based on the single-particle method, the Tamor method is based on the dielectric tensor. If agreement is achieved between these two calculations then this is indeed a strong indication for the accuracy of our code. Note that we vary \( \theta \) from large values (89°) when the radiation is almost completely linearly polarized to small values when circular polarization is dominant. In the former case \( \alpha_\pm \) as calculated by us mostly depends on the value of \( q \), while in the latter case \( \alpha_\pm \) is mostly dominated by \( v \). Therefore we can test here the accuracy of the calculations for both \( q \) and \( v \). The fact that these two methods agree to better than 0.02 % indicates that our numerical code is very accurate. The values of \( \alpha_\pm \) as presented here in all tables are in units of \( \omega_p^2/\omega_e c \). This way the results are independent of the magnetic field and the number density of electrons.

The main use of the relation between \( \alpha_\pm \) and \( \kappa \), \( q \) and \( v \) is that analytic approximations of the cyclotron opacities \( \alpha_\pm \) (for which there are many) can be used to obtain analytic approximations for the opacities in the Stokes formalism (for which, as far as we are aware, there are none). Such
Table 4. Comparison between absorption coefficients $\alpha_\pm$ obtained using the Tamor method and using the Stokes method together with Eq. (6.20)

<table>
<thead>
<tr>
<th>$kT$ (keV)</th>
<th>$\theta$ (°)</th>
<th>$s$</th>
<th>$\alpha_-$</th>
<th>$\alpha_+$</th>
<th>$\alpha_-$</th>
<th>$\alpha_+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>89</td>
<td>5</td>
<td>9.57E-04</td>
<td>9.58E-04</td>
<td>3.43E-05</td>
<td>3.43E-05</td>
</tr>
<tr>
<td>20</td>
<td>89</td>
<td>10</td>
<td>5.08E-06</td>
<td>5.08E-06</td>
<td>2.48E-07</td>
<td>2.48E-07</td>
</tr>
<tr>
<td>20</td>
<td>89</td>
<td>15</td>
<td>9.99E-08</td>
<td>1.00E-07</td>
<td>7.71E-09</td>
<td>7.72E-09</td>
</tr>
<tr>
<td>20</td>
<td>80</td>
<td>5</td>
<td>7.29E-04</td>
<td>7.29E-04</td>
<td>1.30E-04</td>
<td>1.30E-04</td>
</tr>
<tr>
<td>20</td>
<td>80</td>
<td>10</td>
<td>3.60E-06</td>
<td>3.60E-06</td>
<td>1.32E-06</td>
<td>1.32E-06</td>
</tr>
<tr>
<td>20</td>
<td>80</td>
<td>15</td>
<td>6.62E-08</td>
<td>6.63E-08</td>
<td>3.03E-08</td>
<td>3.03E-08</td>
</tr>
<tr>
<td>20</td>
<td>60</td>
<td>5</td>
<td>6.50E-04</td>
<td>6.50E-04</td>
<td>1.01E-04</td>
<td>1.01E-04</td>
</tr>
<tr>
<td>20</td>
<td>60</td>
<td>10</td>
<td>2.06E-06</td>
<td>2.07E-06</td>
<td>4.18E-07</td>
<td>4.18E-07</td>
</tr>
<tr>
<td>20</td>
<td>60</td>
<td>15</td>
<td>3.04E-08</td>
<td>3.04E-08</td>
<td>7.38E-09</td>
<td>7.38E-09</td>
</tr>
<tr>
<td>20</td>
<td>30</td>
<td>5</td>
<td>1.23E-04</td>
<td>1.23E-04</td>
<td>1.82E-06</td>
<td>1.82E-06</td>
</tr>
<tr>
<td>20</td>
<td>30</td>
<td>10</td>
<td>8.58E-08</td>
<td>8.59E-08</td>
<td>2.43E-09</td>
<td>2.43E-09</td>
</tr>
<tr>
<td>20</td>
<td>30</td>
<td>15</td>
<td>4.13E-08</td>
<td>4.13E-08</td>
<td>1.60E-11</td>
<td>1.60E-11</td>
</tr>
<tr>
<td>50</td>
<td>89</td>
<td>5</td>
<td>5.14E-03</td>
<td>5.14E-03</td>
<td>3.02E-04</td>
<td>3.02E-04</td>
</tr>
<tr>
<td>50</td>
<td>89</td>
<td>10</td>
<td>1.52E-04</td>
<td>1.52E-04</td>
<td>1.09E-05</td>
<td>1.09E-05</td>
</tr>
<tr>
<td>50</td>
<td>89</td>
<td>15</td>
<td>1.21E-05</td>
<td>1.22E-05</td>
<td>1.19E-06</td>
<td>1.19E-06</td>
</tr>
<tr>
<td>50</td>
<td>80</td>
<td>5</td>
<td>4.27E-03</td>
<td>4.27E-03</td>
<td>1.03E-03</td>
<td>1.03E-03</td>
</tr>
<tr>
<td>50</td>
<td>80</td>
<td>10</td>
<td>1.09E-04</td>
<td>1.09E-04</td>
<td>4.54E-05</td>
<td>4.54E-05</td>
</tr>
<tr>
<td>50</td>
<td>80</td>
<td>15</td>
<td>8.29E-06</td>
<td>8.29E-06</td>
<td>4.22E-06</td>
<td>4.22E-06</td>
</tr>
<tr>
<td>50</td>
<td>60</td>
<td>5</td>
<td>3.51E-03</td>
<td>3.51E-03</td>
<td>6.70E-04</td>
<td>6.70E-04</td>
</tr>
<tr>
<td>50</td>
<td>60</td>
<td>10</td>
<td>7.83E-05</td>
<td>7.83E-05</td>
<td>2.13E-05</td>
<td>2.13E-05</td>
</tr>
<tr>
<td>50</td>
<td>60</td>
<td>15</td>
<td>5.27E-06</td>
<td>5.27E-06</td>
<td>1.69E-06</td>
<td>1.70E-06</td>
</tr>
<tr>
<td>50</td>
<td>30</td>
<td>5</td>
<td>1.27E-03</td>
<td>1.27E-03</td>
<td>4.16E-05</td>
<td>4.16E-05</td>
</tr>
<tr>
<td>50</td>
<td>30</td>
<td>10</td>
<td>1.13E-05</td>
<td>1.13E-05</td>
<td>6.41E-07</td>
<td>6.42E-07</td>
</tr>
<tr>
<td>50</td>
<td>30</td>
<td>15</td>
<td>4.03E-07</td>
<td>4.03E-07</td>
<td>3.04E-08</td>
<td>3.04E-08</td>
</tr>
</tbody>
</table>
analytic approximations are much faster to compute than the exact expressions 
for the opacities as given in Eq. (6.19) (mainly because of the Bessel functions 
and the required sum and integral). The accuracies of several of these analytic 
expressions have been discussed by Chanmugam et al. (1989). They found that 
the best analytic formulae for the opacities are given by Robinson & Melrose 
(1984; Dulk 1985). These are

$$\alpha_\pm = 2.67 \times 10^{-9} \frac{N \mu^2 (1 - 15/8 \mu)}{B \eta^2 \sin^3 \theta} \frac{3^2 \eta^2 (1 - 1)^{1/2}}{(1 + a_\pm^2)} \times \frac{s_0^{3/2} x^{1/2}}{1} \times$$

$$\left\{ \left[ c_{2\pm} \left( 1 + \frac{0.85 \eta^2}{s_0} \right)^{-1/3} + (1 - n_\pm^2 \beta_0^2)^{1/2} \right]^2 \right.$$ 

$$+ \frac{n_\pm^2 \beta_0^2 \eta^2 \eta_0 \sin^4 \theta}{2(s_0 + \eta^2)} \left\} \times$$

$$(1 - n_\pm^2 \beta_0^2 \cos^2 \theta) \left( 1 + \frac{4.5297 \eta^2}{s_0} \right)^{1/6} Z^{2s_0} \times \exp[-\mu(\gamma_0 - 1)] \quad (6.31)$$

where

$$\gamma_0 = \left[ 1 + \frac{2s}{\mu} \left( 1 + \frac{9x}{2} \right)^{-1/3} \right]^{1/2}; \quad \beta_0 = (1 - \gamma_0^{-2})^{1/2}; \quad x = \frac{s \sin^2 \theta}{\mu}; \quad (6.32)$$

$$n_\pm \approx 1 - \frac{\omega^2}{\eta^2}; \quad \eta_0 = (1 - \beta^2)^{-1/2}; \quad a_+ = -\frac{1}{a_-} = -\frac{[a + (1 + a^2)^{1/2} \cos \theta]}{|\cos \theta|}; \quad (6.33)$$

$$a = \frac{\sin^2 \theta}{2s|\cos \theta|}; \quad s_0 = \gamma_0 (1 - n_\pm^2 \beta_0^2 \cos^2 \theta)s; \quad \beta' = \frac{n_\pm \beta_0 \eta_0 \sin \theta}{(1 - n_\pm^2 \beta_0^2 \cos^2 \theta)^{1/2}}; \quad (6.34)$$

$$s_c = \frac{3x^3}{2}; \quad c_{2\pm} = a_\pm \cos \theta (1 - n_\pm^2 \beta_0^2); \quad Z = \frac{\beta'^{1/\eta_0}}{1 + 1/\eta_0} \quad (6.35)$$

and $\mu = mc^2/kT$. However, these formulae have some limitations. For small 
angles ($\theta < 30^\circ$) they are not very accurate. Also, while the total opacities are 
fairly accurate, the resulting linear and circular polarization can be off by 10%.

This is because the degree of polarization depends on the difference between
\( \alpha_\pm \), and any error in the opacities will therefore show up particularly in the linear and circular polarization.

In the following, we use Eqs. (6.21), (6.29) and (6.30) to calculate the opacities in the Stokes formalism from the Robinson & Melrose (1984) formulae and compare the results to the exact expressions. For this purpose we plot in Fig. 18 the opacity \( \kappa \) and the degree of linear and circular polarization for optically thin radiation for different angles \( \theta \) at \( kT = 10 \) keV. In Fig. 19 we keep \( \theta \) at 90° and vary the temperature. As we can see the overall qualitative behavior of \( \kappa \) and the degree of linear and circular polarization are reflected by the analytic formulae. One also notices that no harmonic structure appears in the analytically derived quantities. This is not surprising as Robinson & Melrose (1984) derived their formulae by averaging over the harmonics, and no harmonic structure can therefore appear. From Fig. 19 we also see the decrease in accuracy of the analytically derived degree of polarization with increasing temperature. This is expected as the radiation is less polarized at high temperatures and Eq. (6.28) becomes less accurate. Another consequence of this equation is that we will get non-zero circular polarization at \( \theta = 90^\circ \). As this is unphysical we set \( \nu \) to zero when \( \theta \geq 88^\circ \). This prevents extreme numerical inaccuracies for the opacity \( \nu \). Overall we can say that the analytic expression can be used at high harmonic numbers (where the harmonic structure is not very important), low temperatures (i.e. \( kT < 50 \) keV) and at angles \( \theta > 30^\circ \). Thereby, one should keep in mind that results from calculations using the analytic formulae should not be overinterpreted. While they may reflect
the qualitative behavior, they may not be accurate enough to support detailed quantitative calculations.

6.3 Accretion columns in AM Her binaries

As we pointed out above, the accretion onto the magnetic white dwarf in an AM Her system occurs onto a fairly small part of the white dwarf's surface, and one can therefore assume that the magnetic field is homogeneous. Such calculations for a three-dimensional shock were done for example by Wu (1989). The structure of the shock in the radial direction was determined by numerically solving the one-dimensional hydrodynamic equations. Because the plasma flow is locked onto the field lines this becomes a one-dimensional fluid flow. Its properties are determined by the continuity, the momentum and the energy equations. These are respectively (Wu et al. 1994)

\[
\frac{d\rho}{dx} + \rho \frac{dv}{dx} = 0
\]

(6.36)

\[
\rho v \frac{dv}{dx} + \frac{dP}{dx} = 0
\]

(6.37)

\[
v \frac{dP}{dx} - \frac{5}{3} \left( \frac{P}{\rho} \right) v \frac{dP}{dx} = -\frac{2}{3} \alpha \rho^2 \left( \frac{P}{\rho} \right)^{1/2} \left\{ 1 + \epsilon_s \left( \frac{P}{P_s} \right)^\alpha \left( \frac{\rho_s}{\rho} \right)^\beta \right\}
\]

(6.38)

where \( \rho \) is the density, \( P \) is the pressure, \( v \) is the velocity and \( x \) is the spatial coordinate. In the energy equation, an ideal gas is assumed with the equation of state \( P = \rho kT/\mu m_p \). For a completely ionized hydrogen gas the mean molecular weight is \( \mu = 1/2 \) and is \( A = 3.9 \times 10^{16} \) in cgs units. If cooling occurs only through bremsstrahlung, then \( \epsilon_s = 0 \). If cyclotron cooling is also important, then \( \epsilon_s > 0 \) and \( \alpha = 2, \beta = 3.85 \) (Wu et al. 1994). To solve this system of equations one typically uses as boundary conditions \( v = 0, T = 0 \) on the surface of the white dwarf. At the post-shock region \( v = -u/4 \) where \( u \) is
Figure 18. Comparison of the opacities $\kappa$, $q$, $v$ calculated from the Robinson \\& Melrose (1984) analytic formulae and the accurate numerical opacities for $kT = 10$ keV and different angles $\theta$. (a) $\theta = 20^\circ$, (b) $\theta = 40^\circ$, (c) $\theta = 60^\circ$, (d) $\theta = 80^\circ$. The figure shows $\kappa$ from the numerical solution (solid) and from the analytic formulae (dash dotted), circular polarization from the numerical solution (dotted) and from the analytic formulae (dash dot dot dotted), linear polarization from the numerical solution (dashed) and from the analytic formulae (long dashed)
Figure 19. Same as Fig. 18 but for $\theta = 90^\circ$ and different temperatures. (a) $kT = 10$ keV, (b) $kT = 20$ keV, (c) $kT = 50$ keV, (d) $kT = 100$ keV
the free-fall velocity. Wu et al. (1994) derived a closed integral form solution of these 1D hydrodynamic equations that can be solved easily. They obtained this formula by introducing dimensionless quantities \( \tau = -v/u \) and \( \zeta = x/x_s \) where \( x_s \) is the shock height. Then the above hydrodynamic equations become

\[
\frac{d\zeta}{d\tau} = \frac{u^2}{2x_s A \rho_a} \frac{\tau^2(5 - 8\tau)}{\sqrt{\tau(1 - \tau)}} K(\tau)
\]

(6.39)

\[
K(\tau) = \left[ 1 + \frac{4^{\alpha + \beta}}{3^\alpha} \epsilon_s (1 - \tau)^{\alpha \beta} \right]^{-1}
\]

(6.40)

where \( \rho_a = \rho_a/4 = \dot{m}/u \) is the pre-shock density, \( \rho_s \) is the post-shock density and \( \dot{m} \) is the specific accretion rate. Note that the dependence on the kind of cooling is contained entirely in the function \( K \). By integrating this equation one gets for the shock height and the structure of the shock

\[
x_s = \frac{u^2}{2A \rho_a} \int_0^{1/4} d\tau \frac{\tau^2(5 - 8\tau)}{\sqrt{\tau(1 - \tau)}} K(\tau)
\]

(6.41)

\[
x_s = 1.55 \times 10^8 \text{ cm} \left(\frac{\dot{m}}{4 \text{ g cm}^{-2} \text{ s}^{-1}}\right)^{-1} \left(\frac{M_{\text{WD}}}{0.5 \text{ M}_\odot}\right)^{3/2} \left(\frac{R_{\text{WD}}}{10^9 \text{ cm}}\right)^{-3/2} \times
\]

\[
\int_0^{1/2} d\tau \frac{\tau^2(5 - 8\tau)}{\sqrt{\tau(1 - \tau)}} K(\tau)
\]

(6.42)

\[
\zeta(\tau) = \frac{u^2}{2x_s A \rho_a} \int_0^\tau d\tau' \frac{\tau'^2(5 - 8\tau')}{\sqrt{\tau'(1 - \tau')}} K(\tau')
\]

(6.43)

The structure of the shock is only a function of the pressure and density dependence of the cooling (i.e. the type of cooling). Pressure and density are scaled by the specific accretion rate and the free-fall velocity. We have used this analytic formula to calculate the structure of the shock. In this case the accretion rate \( \dot{m} = 0.78 \text{ g cm}^{-2} \text{ s}^{-1} \) is uniform across the accretion region which has a cylindrical shape of radius \( 10^7 \text{ cm} \). This corresponds to an accretion rate of \( \dot{M} = 4 \times 10^{-12} \text{ M}_\odot \text{ yr}^{-1} \). The white dwarf has a radius of \( 9.24 \times 10^8 \text{ cm} \) and
a mass of $0.5 \ M_\odot$. The cooling occurs through bremsstrahlung. This gives an electron number density at the shock of $5.9 \times 10^{15} \ \text{cm}^{-3}$, a temperature of $kT = 11.8 \ \text{keV}$ and a shock height of $4.5 \times 10^7 \ \text{cm}$. The magnetic field strength is $30 \ \text{MG}$. We have chosen these values in order to compare the results to the calculation of Wu (1989). In his dissertation the relevant figures are Fig. 12a for the total intensity and Fig. 12c for the circular polarization. No figure is presented in Wu (1989) for the linear polarization in the case of uniform accretion. However, the qualitative behavior can be seen from Fig. 12d, which is for an axisymmetric accretion rate.

For the calculation of the spectra of accretion shocks we start out using the Robinson & Melrose (1984) formulae under the assumption of large Faraday rotation and pulsation and calculating $I_\pm$. In Fig. 20a we show the average intensity. We calculate this by using the equation

$$I = \frac{\sum I_i \delta A_i}{\sum \delta A_i}$$

where $I_i$ is the specific intensity of the $i$th ray and $\delta A_i$ is the area corresponding to this ray. As the projected area can change depending on the viewing direction we keep the area $A = \sum \delta A_i$ over which we average constant for all viewing directions. To get the flux observed from a distance $d$ from the system, we only have to multiply $I$ by the solid angle $\Omega = A/4\pi d^2$. The cyclotron emission is optically thick at small harmonic numbers $s$ and the intensity therefore follows approximately the Rayleigh-Jeans law with $I \sim s^2$. Any deviation is due to the inhomogeneity of the temperature in the shock. The difference in intensity for different viewing directions at the same wavelength is caused also by the inhomogeneity of the temperature and by a projection effect. The maximum
intensity is reached when the emission starts to become optically thin as the
absorption coefficient decreases toward larger frequencies. In Fig. 20b we show
the resulting degree of circular polarization, while Fig. 20c shows the degree of
linear polarization. The degree of linear polarization is approximately zero at
small $s$ when the radiation is essentially coming from optically thick regions.
As $s$ increases and the shock becomes optically thin, the degree of linear po­
larization rises rapidly. For $\theta < 90^\circ$ the linear polarization reaches a maximum
and then decreases slightly as $s$ increases further. This effect was not pointed
out by Wu (1989) as his calculations only go up to harmonic numbers of $s = 10$.
Overall, a good agreement between our calculations and those of Wu (1989) is
achieved.

Next we calculate the spectrum of the accretion shock using the Stokes
formalism. But instead of using the accurate numerical opacities, we compute
$k$, $q$ and $v$ from the analytic formulae for $\alpha_\pm$ by Robinson & Melrose (1984)
using Eqs. (6.21), (6.29) and (6.30). As the condition of large Faraday rotation
and pulsation is fulfilled here, we should obtain the same result as the one in the
previous paragraph when we used the $\pm$-formalism. However, as Eqs. (6.29)
and (6.30) are only valid approximately we cannot expect perfect agreement
between these two methods of calculating the spectra. Still, we find that the
resulting intensity and linear and circular polarization differ by less than 0.1 %
from those calculated as described in the previous paragraph. Only for $\theta > 89^\circ$
does the circular polarization differ significantly, because we set the opacity $v$
to zero. But this is of no importance as the circular polarization goes to zero
for $\theta \to 90^\circ$. The good agreement is expected, as the condition of large Faraday
Figure 20. Theoretical spectrum of the cyclotron emission of an AM Her system. The solid line shows the spectrum obtained by using the accurate numerical opacities in the Stokes formalism. The dotted line shows the spectrum obtained using the Stokes formalism but calculating \( \kappa, q, v \) using the Robinson & Melrose (1984) analytic formulae. For the cases shown in the figure, this gives about the same results as when the transfer is solved in the \( \pm \)-formalism using the Robinson & Melrose (1984) analytic formulae. (a) average intensity \( I \) (in erg cm\(^{-1}\) s\(^{-1}\) Hz\(^{-1}\) sterad\(^{-1}\)) versus the harmonic number \( s \). Curves 1, 2, 3, 4 correspond to \( \theta = 30^\circ, 45^\circ, 60^\circ, 75^\circ \) respectively. The intensity is averaged over an area of \( 4 \times 10^{14} \) cm\(^2\). (b) degree of circular polarization. The curves are for the same angles \( \theta \) as in (a). (c) degree of linear polarization. Curves 1, 2, 3, 4 are for \( \theta = 90^\circ, 89^\circ, 88^\circ, 87^\circ \) respectively.
pulsation and rotation (see Appendix A) is fulfilled. Any deviation between the two computations should be the result of the inaccuracy of the assumption underlying the use of the Robinson & Melrose (1984) formulae for the Stokes formalism. But at the temperatures in this shock Eq. (6.28) is approximately valid.

Finally, we perform the computations for this shock using the correct cyclotron opacities as calculated from Eqs. (6.8) to (6.11). As the calculation of the opacities take up considerable computing time we do not calculate $\kappa$, $q$ and $v$ directly from these equations at every point. Instead we calculate a table of opacities for each used harmonic number and angle $\theta$ but only a small number of temperatures. We use temperatures from 0 keV to 20 keV in 0.5 keV steps and interpolate for the temperatures occurring in the shock. The values of the opacities stored in the array are given in units of $\omega_p^2/\omega_c c$. Therefore, these values are independent of the number density of electrons and the magnetic field. Within the shock itself the number density varies while the magnetic field is constant. By calculating $\omega_p^2/\omega_c c$ at each point in the shock we can then recover the correct opacities. We present the results of our calculations in Figs. 20a to 20c, as well. One can see that the dependence of the intensity on $\theta$ and $s$ is about identical to when we use the Robinson & Melrose (1984) formulae. The dependence of the linear and circular polarization is also nearly the same. In particular, for $\theta$ close to 90° the degree of linear polarization has a maximum at some harmonic number and then decreases slightly as $s$ increases further. This is therefore a real effect and is not caused by the use of the only approximately correct Robinson & Melrose (1984) formulae. However, one
can also notice differences between the two cases. As the angle $\theta$ increases the polarization behaves less smoothly. Particularly at $\theta$ close to $90^\circ$ there are some harmonic numbers when the degree of linear and circular polarization is not zero as expected when the total intensity follows approximately the Rayleigh-Jeans law. This is an effect of the strong harmonic structure of the opacities at small $s$ and low temperature. At the top of the shock, where the temperature is maximal, the opacity $\kappa$ does not have a strong harmonic structure, $\kappa$ is large, this region of the shock is optically thick and light is unpolarized. Because the temperature there is maximal most of the light is coming from this region, overall the intensity is in the Rayleigh-Jeans region and light is unpolarized. However, at an angle close to $90^\circ$ the observer also can look into regions where the temperature is smaller (at the bottom of the shock the temperature is approximately zero). There $\kappa$ has a strong harmonic structure and in between the peaks the radiation is optically thin. This light therefore is polarized. Though it does not contribute much to the overall radiation because of the lower temperature, it is sufficient to cause a degree of linear polarization of a few percent.

6.4 3D effects in an extended polar cap

In many calculations of the cyclotron emission the calculations are simplified by regarding the emission region as a two-dimensional region and the length of the ray through the shock is assumed to be constant or it is calculated from trigonometric considerations. An example of the former are the calculations of Piirola et al. (1993), while the latter was done by Chanmugam & Frank (1987) and Frank & Chanmugam (1990). However, both kinds of
calculations miss an important effect, which results in an underestimate of the emitted radiation. We will describe this effect in the following.

As we pointed out in the introduction to this chapter, polarization has only been observed in two intermediate polars so far (and one of those only two years ago). This could mean that these systems have weak magnetic fields in comparison to AM Her systems. Nevertheless, an evolution from IPs to AM Her systems has been suggested, and this would require that IPs have magnetic fields of about the same magnitude as AM Hers (Chanmugam & Ray 1984; King et al. 1985). Therefore, it has been suggested that the lack of observed polarized radiation from IPs may mean that the polarization is not observed only because of dilution effects. The x-ray emission from IPs shows that accretion in these systems must occur onto extended regions. As the magnetic field across such regions cannot be regarded as homogeneous anymore, this may be a possible source of dilution of the polarized radiation. This was the motivation of the calculations by Chanmugam & Frank (1987) who computed the emission from an extended polar cap in order to investigate the effect of the dipole field on the observed degree of polarization.

In Fig. 21 we show the basic geometry of a polar cap. The polar cap extends in latitude by $\delta_{\text{max}}$. In the calculations by Chanmugam & Frank (1987), the shock height was considered small in comparison to the radius of the white dwarf. Therefore the height dependence of the dipole field was neglected and the dipole field was given as $\mathbf{B} = (B_0/2)(3 \sin \delta \cos \delta \sin \phi, 3 \sin \delta \cos \delta \cos \phi, 3 \cos^2 \delta - 1)$. They proceeded by calculating the emitted intensity of all rays which originated from the emission region on the surface of the white dwarf.
Figure 21: Geometry of a polar cap. It has an extend in latitude of $\delta_{\text{max}}$.

We repeat their calculations and show the resulting flux and polarization as a function of inclination for two different polar caps, one with a maximal extent of $\delta_{\text{max}} = 60^\circ$ (Fig. 22a) and the other with $\delta_{\text{max}} = 10^\circ$ (Fig. 22b). The polar cap has $kT = 20$ keV, $N = 10^{15}$ cm$^{-3}$, $B_0 = 7.5 \times 10^6$ G and a height of $10^7$ cm (the radius of the white dwarf is $5 \times 10^8$ cm). The transfer is solved at a wavelength of 7500 Å, which corresponds to a harmonic number of $s = 19$, and one can consider the shock to be optically thin to cyclotron radiation. While we usually store the input data on a cartesian grid, we use here a grid based on a spherical coordinate system to store the input data. This requires only minor changes in the program.

Figure 23 shows the geometry at a large inclination. As one can see the path length of a ray which passes through the shock just above the surface...
Figure 22. Cyclotron emission from an extended polar cap where radiation originates from the surface of the star and where all rays have been included. (a) for $\delta_{\text{max}} = 60^\circ$ and (b) for $\delta_{\text{max}} = 10^\circ$. For the case that only rays originating from the surface of the star are included, we give the average intensity (long dashed), the degree of circular polarization (dash dot dot dotted) and the degree of linear polarization (dash dotted). For the case that all rays are included, we give the average intensity (solid), the degree of circular polarization (dashed) and the degree of linear polarization (dotted). The intensity is averaged over an area of $10^{18}$ cm$^2$. 
of the white dwarf is twice as long as that of a ray an infinitesimal distance lower, as this latter ray is stopped by the surface of the white dwarf. In the case of optically thin radiation, the emitted intensity of the former ray is then twice as high as that of the latter ray. Therefore it is important to also include those rays in the calculation which do not directly originate from the surface of the white dwarf but only skip through the shock just above the surface. When the rays pass over the magnetic pole of the white dwarf, it is particularly important to include these rays. The cyclotron opacity is largest when the magnetic field is perpendicular to the line of sight and the rays passing just above the surface of the white dwarf will therefore carry most of the energy. We illustrate this in Figs. 24a and b, which show the spatially resolved emission of the polar cap for $\delta_{\text{max}} = 60^\circ$ when $\theta = 90^\circ$. In Fig. 24a we do not include rays that only skip through the shock without hitting the white dwarf surface, while these rays are included in Fig. 24b. One can see that most of the radiation is coming from a small strip along the edge of the white dwarf rim, as it is in this strip where the inclination between the magnetic field and the line of sight is largest. Furthermore, one can see by comparing Figs. 24a and b, which are to the same scale, that a large fraction of the emission is missed in Fig. 24a.

The new spectra, that include all rays, are also shown in Figs. 22a and b, so that they can be compared easily to the spectra that do not include all rays. While these two spectra are essentially identical for small $\theta$, differences occur for $\theta \rightarrow 90^\circ$. At large inclinations, the effect of including the correct path length on the rim is to increase the intensity, decrease the circular polarization and increase the linear polarization. This is most easily understood by looking
Figure 23. Rays emitted from an extended 3-dimensional polar cap. The thick line represents the surface of the white dwarf, and the thin line marks the upper edge of the shock. The ray A is emitted from the surface of the white dwarf, while the ray B is a small distance above ray A and does not touch the white dwarf surface. If the distance between ray A and B is infinitesimally small the intensity of ray B will be twice that of ray A (if the emission region is optically thin).

at the extreme case of $\theta = 90^\circ$. As we saw by comparing Figs. 24a and b more radiation is originating from the rim of the white dwarf if all rays are included than if only rays coming from the surface of the white dwarf are included (about a factor 2.5 more for $\delta_{\text{max}} = 60^\circ$). But at the rim the magnetic field is perpendicular to the line of sight and the light from this region is linearly polarized. This explains the increase of linear polarization (by a factor of $\sim 2$ for $\delta_{\text{max}} = 60^\circ$) and the decrease of circular polarization (by $\sim 10\%$).

6.5 The intermediate polar RE 0751+14

As we mentioned in the introduction, Pirola et al. (1993) estimated that RE 0751+14 has a magnetic field of 8-18 MG. They obtained this result by multiplying the maximum intensity with the observed degree of polarization and plotting this against the frequency. They then fitted the Wickramasinghe
Figure 24. Images of the emission region of an extended polar cap with $\delta_{\text{max}} = 60^\circ$. In (a) only rays from the surface are included, while in (b) all rays have been included. Both images are on the same linear grey scale.

& Meggitt (1985) cyclotron emission model to this plot. The intensity from any cyclotron source has a maximum at the wavelength where the system has the transition from being optically thin to being optically thick. The wavelength where this maximum occurs is sensitive to the parameters (particularly the magnetic field strength) that describe the shock region. Unfortunately, the maximum is at longer wavelengths than observed, and one can only fit the light curve at wavelengths where the system is optically thin. In addition to obtaining an estimate of the magnetic field strength Pirola et al. (1993) also fitted the observed light curve by using the extended ribbon model by Wickramasinghe et al. (1991) and using a radial magnetic field. In the extended ribbon model the accretion does not occur directly onto the magnetic pole but instead onto ribbon-like regions on the white dwarf. We show the geometric
structure of the model in Fig. 25. Two accretion poles are located at mean magnetic colatitudes of $\beta_1 = 45^\circ$ and $\beta_2 = 125^\circ$. Each pole has an extent of 40° in latitude and has a width of 30°. The intensity from each pole is weighted by a factor. A weighting factor of 0.5 is used for the upper pole, while a factor of 1.0 is used for the lower pole. This implies that the accretion rates onto the two poles are not identical. Instead more matter accretes onto the lower pole. There are various ways by which the symmetry can be broken. One is that the accretion disk, which feeds the accretion streams onto the poles is not axisymmetric. The difference in weighting factors for the two poles is necessary as the ribbon model cannot reproduce the light curve otherwise. The magnetic field at the poles is assumed to be radial. The system is seen at an inclination of 60°. The basic features of the light curves as calculated by Piirola et al. (1993) matched the observations.

In the following, we reproduce the light curve of RE 0751+14 using our code. As in the case of extended polar caps we store the input data on a spherical grid. We use an isothermal shock of $kT = 20$ keV and a magnetic field of $B = 10$ MG. The height of the shock is $10^7$ cm with a radius of the white dwarf of $10^9$ cm. We choose an electron number density of $\sim 10^{14}$ cm$^{-3}$ such that the plasma parameter is $\Lambda = 10^6$. As in the calculations of the radiation from AM Her systems we tabulate the exact opacities $\kappa$, $q$ and $v$ but this time we interpolate for the angle $\theta$. In these calculations we also include the free-free opacity in addition to the cyclotron opacity. This was also done in the models on which the work by Piirola et al. (1993) is based. The free-free opacity is given by (Wickramasinghe & Meggitt 1985; there is a misprint in
Figure 25. The model used to calculate the light curve of RE 0751+14. When the accretion pole is hidden from the view its region is dotted. The upper pole produces positive circular polarization while the lower pole produces negative circular polarization.

their expression for $v_H$)

$$
k_H = \frac{2s^4 + 2s^2 - 3s^2 \sin^2 \theta + \sin^2 \theta \left( \frac{\omega_p}{\omega} \right)^2 \nu_c}{2(s^2 - 1)^2} \frac{\nu_c}{c} \quad (6.45)$$

$$q_H = \frac{\sin^2 \theta (1 - 3s^2)}{2(s^2 - 1)^2} \left( \frac{\omega_p}{\omega} \right)^2 \nu_c \frac{\nu_c}{c} \quad (6.46)$$

$$u_H = 0 \quad (6.47)$$

$$v_H = -\frac{2s^3 \cos \theta}{(s^2 - 1)^2} \left( \frac{\omega_p}{\omega} \right)^2 \nu_c \frac{\nu_c}{c} \quad (6.48)$$

where the collision frequency is approximately

$$\nu_c = 3.63NT^{-3/2} \ln(2.95 \times 10^{11}T/\omega) \quad (6.49)$$

with all quantities in cgs units. The total opacities are obtained by adding the cyclotron and the free-free opacities. In Fig. 26 we show our results. Like Pirola et al. (1993) we can reproduce the main observed features. These
include the asymmetry of the light curve and the presence of both positive and negative circular polarization during extended times. Furthermore, the peak in the intensity occurs just before the sign of the circular polarization changes from $-$ to $+$, and the peak in the linear polarization occurs just during this sign change. The position angle varies smoothly between values less than $180^\circ$ apart.

Finally, one can try to fit the modeled maximum intensity to the observed maximum intensity. However, the observed light is a mixture of light from the white dwarf, the secondary, a possible accretion disk, and the cyclotron radiation which one actually models. To make a fit one therefore has to first extract the cyclotron part of the light from the observed light curve. As only this light is polarized one can get it by using the observed degree of circular polarization. Piirola et al. (1993) do this by multiplying the observed maximum flux $F_\omega$ by the observed maximum degree of circular polarization $P_c$ at each observed frequency $\omega$. However, they thereby do not take into account that the cyclotron radiation from the system is not completely polarized and that its degree of polarization may vary with frequency. To actually get the cyclotron radiation one still needs to divide $P_c F_\omega$ by the calculated maximum degree of circular polarization $P_{c \text{ mod}}$ of the model. Then one can fit the model light curve $F_{\omega \text{ mod}}$ to the quantity $P_c F_\omega / P_{c \text{ mod}}$. Alternatively one can fit $P_c F_\omega$ to $P_{c \text{ mod}} F_{\omega \text{ mod}}$. We adopt the latter approach because we have to extract the observed quantity $P_c F_\omega$ from Fig. 3 of Piirola et al. (1993). In Fig. 27 we show our results. There we also show the results when we do not consider $P_{c \text{ mod}}$ the way we described it above. As the degree of circular polarization for the three
Figure 26. Modeled light curve of RE 0751+14 over one period. The different figures give (a) the intensity, (b) the degree of circular polarization, (c) the position angle, (d) the linear polarization. In each figure the dashed curve is from the upper pole, the dash dotted curve is from the lower pole and the solid curve is the combined light curve of both poles.
wavelengths is nearly constant the two curves are almost identical. However, it will be important to include the polarization of the model when observations near the peak intensity are available, as the circular polarization drops sharply near this peak. We use models with $kT = 20$ keV and $kT = 10$ keV. In both cases we calculate the spectra for magnetic fields of $B = 5$, 10 and 15 MG. One can see that at $kT = 20$ keV we can set firm limits on the magnetic field strength. We get $5 < B < 15$ MG. For $kT = 10$ keV we get $5 < B \leq 15$ MG as the $B = 15$ MG is still marginally possible considering the size of the error bars. In the case of $B = 5$ MG the free-free opacity is larger than the cyclotron opacity at these wavelengths. As $\kappa_{\text{ff}} \approx (\omega_p/\omega)^2 (\nu_e/c)$ for $s \gg 1$, which is the case here, the frequency dependence of $\kappa_{\text{ff}}$ is much smaller than that of the cyclotron opacity and we obtain a spectrum that is too flat in comparison to the observed one. Overall we therefore conclude from our calculations that the magnetic field in RE 0751+14 is about 10 MG or larger. This has important implications for fundamental questions regarding the distribution of magnetic fields in magnetic CVs and the evolution of magnetic CVs.
Figure 27. Fit of the model to the observed maximum circularly polarized flux. The observed data points are extracted from Fig. 3 of Pirola et al. (1993) and correspond to the I, the R band and the V band (from left to right). The flux is $F = P_c F_\omega$. The thin curves correspond to the modeled flux $F = F^{\text{mod}}$ while the thick curves correspond to the flux $F = P_c^{\text{mod}} F^{\text{mod}}$. All curves have been normalized to the observed flux in the I band. The frequency is given in terms of the harmonic number $s$ with a magnetic field of 10 MG. In the upper figure we have $kT = 20$ keV and in the lower figure we have $kT = 10$ keV. In both figures we have magnetic field strengths of 5 MG (dashed), 10 MG (solid), 15 MG (dotted).
CHAPTER 7

NLTE RADIATIVE TRANSFER

7.1 Introduction

In the case of non-local thermal equilibrium (NLTE), the source function, on which the radiation field depends, is itself a function of the radiation field. The transfer equation then becomes an integral-differential equation. The calculation of the source function then becomes very difficult. To solve this problem, one generally resorts to iterative methods, since the size of the problem can easily prohibit the storage of all the relevant quantities. The classical method is the $\Lambda$-iteration (Mihalas 1978). However, the convergence rate can be slow in interesting problems. For that reason, different ways of accelerating the convergence have been developed. Operator splitting was first used by Cannon (1973a, b) for radiative transfer problems. It was then further developed by, among others, Scharmer (1981) and Werner & Husfeld (1985), the latter introducing the term accelerated $\Lambda$-iteration (ALI). Klein et al. (1989) proposed a new double-splitting method to get a faster convergence. Recently, Turek (1993) proposed a conjugate gradient-like method to obtain the source function, and Turek & Wehrse (1993) applied it to astrophysical problems.

In this chapter we proceed as follows. First we give a short description of the theory of radiative transfer and acceleration methods as far as it is necessary here. To solve the NLTE problem we have developed a program based on the so-called short characteristic method, which we have adapted to
the MasPar MP-1, and we give a description of this method together with a
discussion of the timing of the necessary computations. Then we apply this
program to two cases. The first involves the situation of Dirichlet boundary
conditions, in which the radiation entering the computational domain is known.
In the second case we have periodic boundary conditions in two dimensions
and Dirichlet boundary conditions in the third dimension. This situation is
important for calculations, which take the 3D structure of stellar atmospheres
into account. For both cases we discuss the convergence properties of different
acceleration methods.

7.2 Radiative transfer

To calculate the spectrum of an object in three dimensions, one must
solve the radiative transfer equation (Mihalas 1978)

\[ \vec{n} \nabla I_\nu(\vec{r}, \vec{n}) = \frac{d}{ds} I_\nu(\vec{r}, \vec{n}) = \chi_\nu(\vec{r}, \vec{n})(S_\nu(\vec{r}, \vec{n}) - I_\nu(\vec{r}, \vec{n})) \]  

(7.1)

where \( \chi \) is the opacity, \( S \) is the source function, \( I \) is the specific intensity, and
no time dependence is considered. The vector \( \vec{n} \) gives the direction of the ray,
\( \nu \) is the frequency, and \( \vec{r} \) gives the position. The parameter \( s \) is the length
along the ray propagating in the direction \( \vec{n} \). The transfer equation can easily
be integrated along the line of sight, and thus one obtains the formal solution

\[ I_\nu(s) = I_\nu(0) \exp(-\tau_\nu(s)) + \int_0^{\tau_\nu(s)} d\tau'_\nu S_\nu(\tau'_\nu) \exp(\tau'_\nu - \tau_\nu(s)) \]  

(7.2)

\[ \tau_\nu(s) = \int_0^s \chi_\nu(s')ds' \]  

(7.3)

where \( \tau \) is the optical depth along the line of sight. In stellar atmospheres the
radius is normally the only spatial variable, and one needs to consider only one
angle for the vector $\mathbf{n}$. However, if the object has no spherical symmetry, all quantities depend in general on three spatial variables, and $\mathbf{n}$ is a function of two angles. As the opacity is normally also frequency dependent, the specific intensity becomes a function of six variables.

For a two-level atom in statistical equilibrium, the source function can be written as

$$S = (1 - \epsilon)J + \epsilon B$$

(7.4)

where $B$ is the Planck function, $J$ is the mean intensity, and $\epsilon$ is the thermalization parameter. The latter represents the probability that a photon is absorbed and therefore converted into thermal energy. The mean intensity for the two-level atom becomes

$$J = \frac{1}{4\pi} \int d\Omega \int d\nu \Phi(\mathbf{n}) I(\mathbf{n}) \,.$$  

(7.5)

The function $\Phi$ is the absorption profile. Clearly the source function is not only dependent on local quantities when $\epsilon < 1$, but also depends on the radiation field, which the transfer equation makes a strongly non-local quantity. The specific intensity and therefore $J$ depend on the source function, and one can write $J = \Lambda[S]$ where the $\Lambda$-operator (in the case of the two-level atom a linear operator) is acting on $S$. After discretization of the problem onto a spatial grid, one obtains a system of linear equations $J = \Lambda S$ where $\Lambda$ is now a matrix. In principle, one can directly solve for the source function by matrix inversion

$$S = [1 - (1 - \epsilon)\Lambda]^{-1}\epsilon B$$

(7.6)

where $1$ is the unity matrix. However, if one has a three-dimensional grid with 64 points in each dimension, the $\Lambda$-matrix has about $6.9 \times 10^{10}$ elements.
Such a matrix cannot be stored on present machines, and an inversion becomes impossible. The problem has to be solved iteratively. The source function of the \((i+1)\)th iteration derives from the \(i\)th source function by

\[
S^{i+1} = (1 - \epsilon)\Lambda S^i + \epsilon B .
\]  

(7.7)

This is the classical \(\Lambda\)-iteration.

As we already pointed out above, the convergence can be accelerated significantly with the ALI method. It is now in common use and has been reviewed by several authors in recent years (Kalkofen 1987; Rybicki 1991; Hubeny 1992). One writes \(\Lambda = \Lambda^* + (\Lambda - \Lambda^*)\), where \(\Lambda^*\) is an approximation to the true \(\Lambda\)-operator. The iteration scheme can then be written as

\[
S^{i+1} - S^i = [1 - (1 - \epsilon)\Lambda^*]^{-1}(S^{FS} - S^i) 
\]  

(7.8)

where \(S^{FS}\) is the source function obtained through a classical \(\Lambda\)-iteration from \(S^i\). This is essentially a Jacobi or block Jacobi iteration depending on the form of the approximate operator (Stoer & Bulirsch 1990). When one uses the diagonal of the \(\Lambda\)-operator, the above matrix inversion becomes a simple division. At large optical depths \(\Lambda \to 1\), and when \(\epsilon \ll 1\) the approximate operator acts as a large amplification factor. Hamann (1985) and Werner & Husfeld (1985) were the first to interpret the operator splitting that way.

Now arises the question of how to construct the approximate operator. There have been a number of proposals, and we refer to the review articles for references. However, it was Olson et al. (1986) who showed mathematically that a nearly optimal operator is simply the diagonal of the \(\Lambda\)-matrix. One
can find the diagonal by the definition of the $\Lambda$-matrix that

$$
\begin{pmatrix}
J_1 \\
\vdots \\
J_n
\end{pmatrix} =
\begin{pmatrix}
\Lambda_{11} & \cdots & \Lambda_{1n} \\
\vdots & \ddots & \vdots \\
\Lambda_{n1} & \cdots & \Lambda_{nn}
\end{pmatrix}
\begin{pmatrix}
S_1 \\
\vdots \\
S_n
\end{pmatrix}. 
$$

(7.9)

The $i$th column of the $\Lambda$-matrix is therefore the mean intensity, if the source function is zero at all gridpoints except at the $i$th position, where it is one. One can therefore use any method that gives the mean intensity to calculate the $\Lambda$-matrix as well. The mean intensity at the grid point $i$ is the $i$th diagonal element of $\Lambda$.

As Olson & Kunasz (1987) already pointed out, a diagonal $\Lambda^*$ is a local operator as it does not couple different spatial points. Using this local operator has many advantages computationally, as it requires few calculations (no matrix inversion has to be performed) and little memory space, and as it nevertheless performs well in many cases (MacFarlane 1992). However, in the optically thin areas the radiation field is highly non-local, and therefore the off-diagonal elements of the $\Lambda$-matrix are no longer negligible in comparison to the diagonal elements. Therefore, one can expect to improve the convergence by including off-diagonal elements in $\Lambda^*$, which then becomes a non-local operator including spatial coupling. However, what was a simple division in Eq. (7.8) in the case of a diagonal $\Lambda^*$ has now become a matrix inversion. In the one-dimensional case, including the nearest neighbors results in a tridiagonal matrix (Olson & Kunasz 1987; Werner 1989). In the two-dimensional case, one obtains a matrix with nine non-zero elements per row, which was used by Steiner (1990). But in the 3D case, a matrix inversion becomes quite difficult as we get twenty-seven non-zero elements per row with many zero elements between
non-zero ones in each row. When we therefore include off-diagonal elements, we will solve Eq. (7.8) not by directly inverting the matrix, but by solving the corresponding linear system of equations iteratively using the Jacobi method. When we have some approximation $S'$ for the source function, then the $j$th step of this iteration can be written as

$$
\tilde{x}^{j+1} = \tilde{x}^j + [1 - (1 - \epsilon)\Lambda^D]^{-1} \{(1 - \epsilon)\Lambda^*\tilde{x}^j + \tilde{b} - \tilde{x}^j\}
$$

(7.10)

where $\tilde{x}^j = S^j - S'$, $\tilde{b} = S^{FS} - S'$, and $S^{FS}$ is the source function we obtain from $S'$ using a classical $\Lambda$-iteration. The matrix $\Lambda^*$ is an approximation to the $\Lambda$-matrix that includes off-diagonal elements, and $\Lambda^D$ is the diagonal of the $\Lambda$-matrix. While we avoid the complicated matrix inversion, we need to make several iteration steps before we get the vector $\tilde{x}$ with a sufficient accuracy. But as the elements of the diagonal of the $\Lambda$-matrix are always the largest elements per row and column, we can expect to get a sufficiently accurate solution quite rapidly, especially when we use $\Lambda^D$ to get a good initial estimate for the vector $\tilde{x}$.

There are additional ways of accelerating the convergence. We test the Ng method (Ng 1974) that was first introduced by Buchler & Auer (1985) for radiative transfer problems. Furthermore, we use the orthomin method (Vinsome 1976), which was first applied by Klein et al. (1989) to radiative transfer problems. Both can be combined with the ALI method. In addition to the descriptions of these methods in the above-mentioned papers, they have been discussed and compared by Auer (1991), and we therefore will not give a detailed description here. The essential idea behind both methods is to use several previously calculated source functions in order to estimate the correct
source function. The Ng method gets the new estimate by minimizing the residual directly, while the orthomin method minimizes the residual with respect to a set of conjugate vectors. In comparison, the ALI method uses the previously calculated source function and calculates the next source function by a different matrix multiplication. The ALI method therefore represents a different concept from the Ng or orthomin acceleration. The free parameter in these acceleration methods is the number of previously calculated source functions. Here we use three residuals for the Ng method and two conjugate vectors for the orthomin method.

Finally, we test the multi-grid method as first introduced to radiative transfer problems by Steiner (1991). Again we will not go into the details here, as they can all be found in depth in his paper, and as there exists a vast literature on this subject (e.g. Hackbusch 1985), but we will describe the general idea of this method. When one has some approximation to a linear equation on a grid of a certain coarseness, then this approximation will have deviations from the exact solution on this grid. There will be variations over the scale of the separation of the grid points (short-period errors) and variations over the entire mesh (long-period errors). Acceleration methods that try to minimize the residual of a linear equation can only reduce the short-period errors efficiently, not the long-period errors. The idea of multi-grid methods is to introduce grids of varying coarseness. On the coarse grids, the large scale variations of the error can be minimized; and on the fine grids, the small-scale variations can be reduced. To get grid quantities from fine grids to coarser grids, one applies a restriction, which we do here by averaging the
quantities. The reverse process from the coarse grid to the fine grid is called the *prolongation*, for which we use linear interpolation. On the coarsest grid one solves the so-called *coarse grid equation*, which has the same structure as Eq. (7.7) for the linear A-operator. To do this we make four steps of the ALI method with a local operator. There exists a vast number of different algorithms on how best to combine the use of grids of varying coarseness. We adopt here the algorithm as described by Steiner (1991).

### 7.3 Solution of the NLTE radiative transfer problem on the MasPar

#### 7.3.1 Short characteristics

When the thermalization parameter is much smaller than one, the source function depends strongly on the radiation field, and we have to use Eqs. (7.7) or (7.8) to get the source function by iteration. After each step we need to calculate the specific intensities for all direction vectors $\vec{n}$ and frequencies $\nu$ on all the grid points. As was pointed out by Castor et al. (1991), we potentially need a very large number $N_\vec{n}$ of vectors $\vec{n}$ for 3D cases to adequately resolve the integrand of Eq. (7.5). Additionally, a large number $N_\nu$ of frequency points can be required when velocities become important due to the Doppler shift (Adam 1990). Therefore, the calculation of the mean intensity will be the most time-consuming part of 3D radiative transfer calculations, and we need a fast numerical method to get the specific intensities on the grid points. If one works on a SIMD machine, one must have a numerical method that uses the PE grid optimally, i.e., most or all processors are active at any given time. This implies that all the information needed by any processor is located on or near to that processor. Otherwise, communication may take up an unreasonable amount of
time in comparison to actual numerical calculations. But the radiation field is intrinsically a non-local quantity through the radiative transfer equation, and therefore the problem of solving this equation is not ideally suited for parallel machines. In the following, we describe a method with which the transfer equation can nevertheless be solved efficiently on a SIMD machine.

A fast method to obtain the specific intensities on a cartesian grid are the short characteristics as described by Kunasz & Auer (1988) for the two-dimensional case, but the concept easily generalizes to three dimensions. In Fig. 28 we show one x-y layer of a 3D cartesian grid and the short characteristic for point p. This is simply the characteristic of the transfer equation in this cell of the grid. If \( \tau \) is the optical depth along this characteristic, then the specific intensity \( I_p \) at the point p is

\[
I_p = I_{p'} \exp(-\tau) + \int_0^\tau d\tau' S(\tau') \exp(\tau' - \tau)
\] (7.11)

where \( I_{p'} \) is the specific intensity at the point \( p' \), from which the short characteristic starts. In order to calculate the integral, one has to decide on the way to interpolate the source function along the short characteristic. We choose linear interpolation here. Furthermore, we have to get \( I_{p'} \) by interpolating the specific intensities from neighboring points. In the simplest case of linear interpolation, these are in Fig. 28 the points k, l, \( \lambda \) and \( \mu \). Because these intensities have to be known first before \( I_{p'} \) can be calculated, we have to go through the grid cell by cell in order to get the intensities at the grid points. It is hereby assumed, that the radiation coming into the grid is known, i.e., that we have Dirichlet boundary conditions.
Figure 28. $x$-$y$ layer of the 3D cartesian grid showing a short characteristic for point $p$.
When one works on the MP-1 with its two-dimensional PE grid, a useful way of storing the data of the cartesian grid with its $N_x \times N_y \times N_z$ points is to map the $x$-$y$ axis of the input grid onto the PE grid and to identify the $z$ axis with the memory of the PEs. Then, ideally, all the specific intensities of one $x$-$y$ layer should be calculated in one step, and in the following we will investigate whether this is possible. First we have to calculate the integral of Eq. (7.11) at every point of a $x$-$y$ layer. As all the data needed to calculate the integral at some point are stored on or nearby the processor that corresponds to that point, we can indeed calculate this integral simultaneously at every point of the $x$-$y$ layer. Necessary data transfer can be done using the X-Net and is therefore fast. Now we have to get the specific intensity at the start of every short characteristic. But as we pointed out above, this leads to a stepping through the input grid, and essentially only one processor is active at a time. To get a more optimal use of the PEs, we require the points of the cartesian input grid to be equidistant in the $x$ and $y$ dimension (the total dimensions of the $x$ and $y$ axis still do not need to be identical). Then all short characteristics in one $x$-$y$ layer start from the same side. Therefore, the geometrical data needed to characterize them scale as $N_x \times N_y$ instead of as $N_x \times N_x \times N_y \times N_z$. Furthermore, in the case that the short characteristics in one $x$-$y$ layer start from the $x$-$y$ layer below, where the specific intensities have already been calculated, we can get the specific intensities in the current $x$-$y$ layer in one step. If the short characteristics start from the $x$-$z$ side or the $y$-$z$ side, we need to sweep through the $x$-$y$ layer in $N_y - 1$ or $N_x - 1$ steps, and $N_x$ or $N_y$ PEs respectively are active at a time.
As we saw in the previous paragraph, we do not have an optimal use of the PE grid when we have to sweep through the \( x \)-\( y \) layer to get the specific intensities at the grid points. We can get around this problem by viewing the intensities in one \( x \)-\( y \) layer as the solution to a system of linear equations. The specific intensity at the point \( p \) can be written as a linear combination of the weighted specific intensities at the points \( k, l, \lambda \), and \( \mu \) plus a constant which corresponds to the integral of Eq. (7.11). As the intensities at the points \( \lambda \) and \( \mu \) are already known, we can simply write \( I_p \) as a linear combination of \( I_k \) and \( I_l \) and a constant, which contains now the integral and the weighted intensities from the neighboring \( x \)-\( y \) layer. This rewriting can be done for every point of the current \( x \)-\( y \) layer. If we use a higher order to interpolate intensities, we simply get more complicated linear equations for the grid points. Overall we can write these linear relations as

\[
A\vec{I} = \vec{I}_0
\]  

(7.12)

where \( \vec{I}_0 \) is a constant vector, and \( \vec{I} \) contains the specific intensities. This vector is unknown prior to the solution of this linear equation except for the components that correspond to points on the boundary in the case of Dirichlet boundary conditions. For the above matrix we have \( A = 1 - W \), where \( W \) is the matrix containing as its elements the weights, which also include the factor \( \exp(-\tau) \), and these elements are therefore much smaller than one at large optical depths between grid points. It is obvious that the diagonal of \( W \) is zero and that all elements of \( W \) are at most one and not negative. Furthermore, there always exists a way to number the grid points such that \( W \) becomes a lower triangular matrix, and the solution to the above algebraic system can
be obtained by forward substitution. But we can also solve this equation iteratively by writing

$$\vec{I}^{i+1} = W \vec{I}^i + \vec{I}_0$$

(7.13)

where $i$ counts the iterations. Using the Jacobi iteration instead of this simple iteration scheme does not make a difference, as the diagonal of $A$ is the unity matrix. The advantage on a SIMD machine is that we can have $N_x \times N_y$ active processors instead of only $N_x$ or $N_y$ when sweeping through the grid. Furthermore, because of the special form of $W$ as a lower triangular matrix, we know that we get the exact solution after $N_x - 1$ or $N_y - 1$ iterations. Even further, we can drop the requirement of having equidistant spacing in the $x$ and $y$ direction and still have the exact solution in $\max(N_x, N_y) - 1$ iterations, as it takes only that many steps to transfer the information from any point on the boundary to any grid point that can be influenced by the conditions at the boundary point. However, more memory space and more arithmetic calculations are then needed, and therefore we keep the grid points equidistant in $x$ and $y$. But the main advantage of using this iteration is the possibility that much fewer steps are required to obtain sufficient accuracy when the elements of $W$ are small. This is the case when the optical depth between grid points is large because of the factor $\exp(-\tau)$. The vector $\vec{n}$ also influences the elements of $W$ through the weights of the interpolation. But most important is the dependence on the angle $\theta$ between $\vec{n}$ and the $x$-$y$ layer. By comparing the intensities obtained by sweeping and by iteration, we find that the iteration has converged when $(\vec{I}^{i+1} - \vec{I})/\vec{I}^{i+1} < 10^{-6}$. Figure 29 shows the required steps until convergence as a function of optical depth $\Delta\tau$. 
Figure 29. Number of steps needed for convergence to get specific intensities in one $x$-$y$ layer for Dirichlet (solid) and periodic boundary conditions (dotted) as a function of the optical depth between grid points for different angles $\theta$. From top to bottom $\theta = 0^\circ, 5^\circ, 15^\circ, 25^\circ, 35^\circ$ between neighboring grid points for different angles $\theta$. The $x$-component of $\vec{n}$ is zero, and an isotropic grid with $N_x = N_y = 64$ points is used. The source function is set to one everywhere. As expected, the convergence is fastest for large $\theta$. For angles larger than $\approx 10^\circ$ the number of iterations is less than $\max(N_x, N_y) - 1$. Furthermore, the number of iterations decreases strongly when the optical depth between neighboring points approaches one, and for depths larger than four we can calculate the specific intensities in a $x$-$y$ layer in one step instead of in $N_x$ or $N_y$ steps.

This iterative solution turns out to be even more useful when we have periodic boundary conditions in $x$ and $y$. In this case, the specific intensities at
the side \( x = x_{\text{min}} \) are equal to the specific intensities at the side \( x = x_{\text{max}} \). The same is the case for the \( y \)-axis, and there are Dirichlet boundary conditions only for rays coming from the top or the bottom of the grid. Such boundary conditions can be important when investigating radiative transfer in stellar atmospheres with inhomogeneities through convection (Nordlund 1991; Steffen 1991) or magnetic fields (Kalkofen et al. 1989). In 2D calculations, in which there are periodic boundary conditions only along one axis, it is possible to calculate the specific intensity at the boundary analytically when using the short characteristic method (Steiner 1990). However, in the 3D case we have for every \( x \)-\( y \) layer \( N_x + N_y - 1 \) unknown specific intensities at the boundary. We can express these intensities by an algebraic equation similar to Eq. (7.12), but with a lower dimension as we have \( N_x + N_y - 1 \) fewer independent intensities. It follows that the matrix \( A \) is no longer triangular and the equation cannot be solved by forward substitution. Therefore, when we solve it iteratively, we cannot expect to find the correct solution after at most \( \max(N_x, N_y) - 1 \) steps, as in the case with Dirichlet boundary conditions. On a SIMD machine one can solve this problem iteratively using again Eq. (7.13) and resetting the periodic boundary conditions after each iteration. As in the case with Dirichlet boundary conditions, we show in Fig. 29 the number of steps required to achieve convergence for different angles \( \theta \) as a function of the optical depth between neighboring grid points. The number of required iterations diverges for \( \theta \to 0 \) and \( \Delta \tau \to 0 \). This is not surprising, as there is no matter and therefore no radiation along the line of sight in this case, but the source function is set to one nevertheless. In order to achieve convergence in less than 128 steps on
this grid, the angle $\theta$ must be greater than $5^\circ$. In a realistic application one can generally fulfill such a restriction.

We now have to discuss how we construct the approximation to the $\Lambda$-matrix for the ALI method. This is done equivalently to Kunasz & Olson (1988). When we want to construct the diagonal element of $\Lambda$ at point $p$ for example (see Fig. 28), we set the source functions to zero at all gridpoints except at the point $p$, where it is set to one. If we use only linear interpolation for the source function, the intensity at point $p'$ is zero, and we get the exact element $\Lambda_{pp}$ by only calculating the integral of Eq. (7.11). Therefore, we can calculate the diagonal of $\Lambda$ for all points in one $x$-$y$ layer simultaneously. However, when we want to construct for example $\Lambda_{pk}$, we have to set the source function to one at point $k$ and zero at all others. Then the intensity $I_{p'}$ is not zero. In order to use the processors optimally, we will set $I_{p'}$ to zero. Therefore, we underestimate the values of the off-diagonal elements. But we can expect this error to be small, especially in the optically thick areas, where the ALI method is most useful.

Finally we have to address the problem of when the iterative scheme has converged to a solution and with what approximation to best start the iteration. As was shown by Olson et al. (1986) and pointed out by Auer (1991) as well, an upper bound of the convergence rate of the classical $\Lambda$-iteration is given by $1 - \epsilon$. Therefore, the solution has converged when $\Delta S/S \ll \epsilon$ where $\Delta S$ is the difference between two source functions obtained by a $\Lambda$-iteration.

To get a reasonably good guess for the initial source function, we have to estimate the mean intensity. We can say that $J \approx (1 - p_e)S$ where $p_e$ is the
escape probability of photons. This probability is given by (Mihalas, 1978)

\[ p_e = \frac{1}{4\pi} \int d\nu d\Omega \exp(\tau_\nu(\vec{n})) \]  

(7.14)

when the effects of scattering are ignored. The optical depth \( \tau_\nu(\vec{n}) \) can be calculated using short characteristics.

7.3.2 Timing

Because the calculations are done simultaneously on many processors, one can expect a significantly different scaling for this code than when the same calculations are done on a conventional scalar or vector machine. This is indeed the case. While the timing on a conventional scalar machine scales as the number of grid points \( N_x \times N_y \times N_z \), the calculation on a SIMD machine scales as \( N_z \) when the processors are used optimally (at least when numerical methods comparable to ours are used). From this we can see the potential power of parallel machines, as we can increase the number of grid points in \( x \) and \( y \) arbitrarily without increasing the computing time, as long as there are sufficient processors and as long as the processors are used optimally. We therefore have to discuss the timing of the numerical methods that we use here in some depth, in order to compare the efficiency of performing the calculation on this SIMD machine with the efficiency when done on a conventional computer. In the following, we will assume that the number of points on the \( x \) and \( y \) axis of the input grid is at most the number of PEs of the \( x \) and \( y \) axis of the PE grid. If this is not the case, then the calculations have to be repeated for all the virtual layers, which we described in Sect. 2.
The most time-consuming part of getting the source function in the case of NLTE is the computation of the mean intensity, because it consists of solving transfer equation for many angles and frequencies, and it has to be repeated for every iteration. We therefore discuss this first. The total timing for it using short characteristics on the MP-1 scales linearly with \( N_\pi \) and \( N_\nu \). We can obtain the timing for the calculation of the specific intensities for a given \( \pi \) and \( \nu \) in the following way. We first have to calculate the integral of Eq. (7.11), for which we need \( A(N_\pi - 1) \) seconds, where the constant is \( A \approx 1.0 \) ms. To calculate the final specific intensities by sweeping one either needs \( A'(N_\pi - 1) \) seconds, if all short characteristics start at a \( x\cdot y \) layer, or \( A'(N_\nu - 1)(N_\pi - 1) \), if they start from a \( x\cdot z \) layer, or \( A'(N_\pi - 1)(N_\nu - 1) \) seconds, if they start from a \( y\cdot z \) layer. For the constant \( B \) we obtain \( B \approx B' \approx 0.25 \) ms (for periodic boundary conditions \( B' \approx 0.45 \) ms). Of course it is also possible that the short characteristics start from different sides in different \( x\cdot y \) layers, if the grid is not equidistant along the \( z \)-axis. But we are interested here in how long a calculation takes on average, and we therefore assume that we have many different viewing directions and a grid with equidistant spacing in all dimensions. The time needed to calculate the mean intensity becomes

\[
T = N_\nu N_\pi \left\{ A + \frac{1}{3} \left[ B + B(N_\pi - 1) + B(N_\nu - 1) \right] \right\} (N_\pi - 1). \tag{7.15}
\]

To simplify the discussion we will assume from now on that we have a grid of \( N \) points in each direction with \( N \gg 1 \). Table 5 lists how the computing time scales for different tasks. For large \( N \) the time scales as \( N^2 \) for the sweeps. When calculating the specific intensities iteratively, the timing depends on the average iteration steps \( \bar{N} \) necessary to get convergence. As discussed before,
### Table 5.
Scaling of the computing time for the different calculations on an isotropic cartesian grid of $N^3$ grid points. The constants are $A \approx 1.0$ ms and $B \approx 0.25$ ms. For details see text.

<table>
<thead>
<tr>
<th>Calculation</th>
<th>Computing time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean intensity by sweeping</td>
<td>$N_\nu N \pi { AN + \frac{2}{3} BN^2 } \approx N_\nu N \pi \frac{2}{3} BN^2$</td>
</tr>
<tr>
<td>Mean intensity by iteration ALI, Ng, orthomin</td>
<td>$N_\nu N \pi { A + \frac{2}{3} B \tilde{N} } \propto N$</td>
</tr>
<tr>
<td>Multi-grid method: $T_0 \propto N^3$</td>
<td>$2.5 T_0$ for $n = 2$</td>
</tr>
<tr>
<td></td>
<td>$2.5 T_0$ for $n = 3$</td>
</tr>
<tr>
<td></td>
<td>$2.5 T_0$ for $n = 4$</td>
</tr>
<tr>
<td>$T_0 \propto N^2$</td>
<td>$3.0 T_0$ for $n = 2$</td>
</tr>
<tr>
<td></td>
<td>$3.25 T_0$ for $n = 3$</td>
</tr>
<tr>
<td></td>
<td>$3.375 T_0$ for $n = 4$</td>
</tr>
<tr>
<td>$T_0 \propto N$</td>
<td>$4.0 T_0$ for $n = 2$</td>
</tr>
<tr>
<td></td>
<td>$5.5 T_0$ for $n = 3$</td>
</tr>
<tr>
<td></td>
<td>$7.0 T_0$ for $n = 4$</td>
</tr>
</tbody>
</table>

This is at most $N$ in the case of Dirichlet boundary conditions, while it can be larger for periodic boundary conditions. However, in the most optimal case $\tilde{N} = 1$, and the computing time scales as $N$.

After we have calculated the mean intensity, we have to get the source function iteratively. To accelerate the convergence, we use ALI and/or Ng or the orthomin acceleration. The computing times for all these methods do not depend on $N_\pi$ and $N_\nu$ and scale simply as $N$. Only in the case of ALI with a non-local operator can the acceleration methods become important for the overall execution time. For one iteration step of Eq. (7.10) we need for a $64^3$ grid 0.15 s with Dirichlet boundary conditions, while it takes 0.45 s in the case of periodic boundary conditions. However, when we use multi-grid methods, where one
uses different numbers $n$ of grids, the calculation of the computing time is more difficult, as one step of the multi-grid method essentially requires the repeated computation of the mean intensity on different grids. The time for one iteration step can therefore be written as $aT_0$, where $T_0$ is the time needed to calculate the mean intensity on a cartesian grid with $N^3$ grid points. As can be seen from Table 5, the factor $a$ depends on how $T_0$ scales with $N$. As $a$ is always larger than one, the convergence rate for a problem using multi-grid methods must be larger by a factor $a$ compared to that of some other acceleration method, or this other acceleration method is faster in total computing time. Clearly, multi-grid methods are of greater advantage on a scalar computer than on an optimally used SIMD machine. To decide whether the multi-grid methods have any advantages over the Ng and orthomin acceleration, we ultimately have to compare their actual performance when applied to radiative transfer problems.

7.4 Discussion – Dirichlet boundary conditions

In the previous section, we have described how to calculate the mean intensity, and how to find the source function by iterative methods. In order to investigate the convergence properties, we assume the ‘grey’ line here, i.e., we only consider one frequency. The 3D cartesian grid we choose has 57 points along each axis and unity spacing between grid points along mesh lines. We use 26 viewing directions. The Planck function is set to one everywhere, and the thermalization parameter is set to $\epsilon = 10^{-3}$, which represents an intermediate value. There is no incoming radiation. The opacity we use is $\chi(x, y, z) = 10^{-2}$, and we interpolate it linearly within the grid cells. This choice of model has the advantage that it contains both optically thin and thick regions. The optical
Figure 30. Solution (solid) and initial guess (dash dotted) for the source function, diagonal of the $\Lambda$-matrix (dotted) and optical depth $\tau$ (dashed)

depth along a ray parallel to the $z$ axis is shown in Fig. 30 as a function of the $z$ coordinate. As one can see, the optical depth increases from zero to about $5 \times 10^5$. Finally we want to point out that we can compare our results with the plane parallel calculations of Auer (1991) because $\chi$ is constant in each $x$-$y$ layer. The two source functions must be identical in the optically thick regions. Only in the optically thin regions do 3D effects become important, and the source function is smaller in the 3D case because of the increased escape probability.
As we described above, one can get an estimate of the source function by calculating escape probabilities using Eq. (7.14). This estimated source function is shown in Fig. 30. This figure (and unless stated otherwise all following figures) shows the source function along the middle axis of the grid parallel to the \( z \) axis. It also shows the 'correct' source function \( S_c \), which was obtained by making many iterations. One can see that the escape probability formulation gives a good approximation to the source function in the effectively optically thick regions where \( \epsilon \tau > 1 \), which causes the photons to be thermalized and the source function to be close to the Planck function. On the other hand, it differs very much from the correct source function in the other regions and only reflects the qualitative behavior of it, as this approximation does not take into account the effects of scattering. But in the following we will always use \( e_B \) as an initial guess for the source function, as we want to demonstrate the convergence behavior of different models, and this value is a lower limit for the source function.

The basic properties of the classical \( \Lambda \)-iteration are shown by Fig. 31a. The source function is still constant at large optical depths, while at small optical depths the convergence has obviously slowed down very much, even though it is still far away from the solution. This can be understood in the following way (Mihalas 1978). The mean free path of photons is about one optical depth. As \( \epsilon \) is the probability of thermalization of photons, they travel by random walk a distance of \( \epsilon^{-1/2} \) before they are thermalized (also called the thermalization length). Therefore, the source function will only become the Planck function at a depth of \( \epsilon^{-1/2} \). As it takes that many steps before
information about the boundary propagates into the interior, the Λ-iteration requires $e^{-1/2}$ steps before it approaches the solution. If there are effectively optically thick regions, one therefore needs to find a different way to obtain the solution.

The solution in the regions where the photons are thermalized can be obtained quickly using the ALI method, which is demonstrated by Fig. 31b when using the diagonal of the Λ-matrix, i.e. a local operator, as an approximation. The diagonal of the Λ-matrix for points along the $z$ axis is shown in Fig. 30. At large depths the Λ-matrix is almost a diagonal matrix. Therefore, we get the final solution there after only a few steps of ALI. However, the ALI method using a diagonal matrix does not give a sufficiently good convergence in the effectively optically thin regions, where the photons are not thermalized. As astrophysically interesting cases will normally contain optically thin regions, we will not achieve sufficiently fast overall convergence. This is illustrated in Fig. 32, which shows the maximum relative error \( |S^i - S_c| / S_c \) of the source function after each iteration step (the maximum relative error for the classical Λ iteration is close to one). A way to improve the convergence is by using a better approximation to the Λ-matrix. In Fig. 31c we show the convergence when we include the 26 off-diagonal elements in $\Lambda^*$ which correspond to nearest neighbors, and use Eq. (7.10) as the iterative scheme to perform one step of ALI. As we can see, the convergence is significantly faster with this non-local operator than when we only use a local operator. In the areas where photons are thermalized, we get the solution after only one step. In the effectively optically thin regions the convergence is also accelerated in
Figure 31. Convergence of a the classical $\Lambda$-iteration, b ALI with local $\Lambda^*$, and c ALI with non-local $\Lambda^*$. The solution is shown by dashed line.
comparison to the local operator. But to decide on the effectiveness of this iteration scheme, we have to look at how many iterations of Eq. (7.10) are needed for one ALI step. Here we limit the maximal number of iterations to at most 128. Then for the first step of ALI, we need this maximal number (though after 20 iterations we already have convergence in the effectively optically thick regions). But for the next step it takes 90 steps to get a relative change smaller than $10^{-4}$ between successive source functions as calculated by Eq. (7.10). In the following, the number of required steps decreases, and after 10 ALI steps we need 43 iterations, which is better than performing a matrix inversion for every ALI step. However, the major disadvantage of this method is the relatively large size of the array required to store the 26 off-diagonal elements. In many cases (e.g. multiple lines) this will exceed the memory space.

The convergence can be significantly accelerated by using the Ng or the orthomin methods of acceleration. Figure 33a shows the convergence of the Ng acceleration, while Fig. 33b shows the convergence of the orthomin acceleration. In both figures we have combined the acceleration methods with ALI using a local operator. Obviously, we get a significantly faster convergence than when only using ALI with a local operator. When comparing these figures one has to keep in mind that one step of the orthomin method essentially requires the calculation of two formal solutions. Then it becomes clear that these two methods are about equally fast in this example. But as is clearly seen in Fig. 32 (and it was noted by Auer 1991 as well), the convergence is much smoother using the orthomin acceleration, and it is therefore preferable.
Figure 32. Maximum relative error of the source function after each iteration step. From top to bottom (at $N = 20$): Classical $\Lambda$-iteration, ALI with local $\Lambda^*$, multi-grid method with $n=2$, Ng acceleration, ALI with non-local $\Lambda^*$, orthomin with local $\Lambda^*$, orthomin with non-local $\Lambda^*$, multi-grid method with $n=3$, multi-grid method with $n=4$. The convergence using orthomin with a non-local $\Lambda^*$ slows down for $N > 12$ because the correction to the source function is smaller than its machine accuracy.
The orthomin method is also tested in combination with ALI using a non-local operator, and Fig. 33c shows the result. This clearly gives the best convergence, but the major drawback is again the large memory space required to store the non-local operator. The number of iterations for one ALI step is as follows. For the first step we again need the maximal number of allowed steps. After that the number of steps rapidly decreases, with 14 iterations for the 10th orthomin step and one iteration for the 16th orthomin step.

Finally, we show the convergence behavior when using multi-grid methods. Figures 34a, b and c show the convergence behavior when \( n = 2, 3, \) and 4, respectively. It can be seen directly that using only two grids does not give a satisfactory convergence. There is little difference between the \( n = 3 \) and \( n = 4 \) case. But as one step using four grids requires more time than one step using three grids (see Table 28), the three grid method is preferable. Overall, these two multi-grid methods are faster than any other method. But because the time needed for one step of the multi-grid method is longer than the time required for one step of orthomin, we have to compare the number of steps needed to achieve a certain convergence. Looking at Fig. 32 and using the numbers from Table 5, we see that the multi-grid method is fastest when the timing scales as \( N^2 \) (when we have to sweep through the \( x\text{-}y \) layers to get the specific intensities), but not if it scales as \( N \) (when we only need one iteration to get the specific intensities in one \( x\text{-}y \) layer). Then the orthogonal minimization using only a local operator still converges faster. Even though there are viewing directions for which the iteration to get the specific intensities requires the maximal number of steps, on average we only need \( \bar{N} = 6 \).
Figure 33. Convergence of a Ng acceleration, b orthomin with local $\Lambda^*$, and c orthomin with non-local $\Lambda^*$. The solution is shown by dashed line.
iterations to get the specific intensities for some viewing direction. Therefore, the timing scales closely to N, and the multi-grid method as implemented here is slower than the acceleration method based on orthogonal minimization. A basic problem of multi-grid methods on SIMD computers is that many processors will be idle when calculations are performed on the coarse grids. A way around this problem was proposed by Frederickson & McBryan (1988). When going from a fine grid to a coarser grid one typically has two choices of coarse grids. In principle, one can therefore do the calculations on both coarse grids and average the results. Thus one should obtain a better estimate of the solution than if one arbitrarily selected one of the coarse grids over the other. On a massively parallel computer one can in principle do the calculations on the second coarse grid (which one would normally not use) on the processors that are otherwise idle. This is the basic idea behind the algorithm proposed by Frederickson & McBryan (1988) and they show that this algorithm is well suited to SIMD computers because the calculations on both coarse grids use the same set of instructions.

7.5 Discussion – Periodic boundary conditions

To illustrate the ability of our program to handle radiative transfer calculations with periodic boundary conditions, we adapt the two-dimensional model of a stellar atmosphere by Steiner (1991) to three dimensions. In Fig. 35 we show the model. The mesh has 49 x 49 x 25 points and the length of the mesh in x and y is two, while it is one in z (in our calculations we do not take advantage of any symmetries). We have periodic boundary conditions for the radiation in x and y. There is no radiation entering into the computational
Figure 34. Convergence of multi-grid methods with a $n=2$, b $n=3$, and c $n=4$. The solution is shown by dashed line.
Figure 35. Part of grey 3D atmosphere with periodic boundary conditions along $x$ and $y$. Isotropic radiation is entering from the bottom. The length of the cube along $x$ and $y$ is 2, while it is unity along $z$. The opacity within the dashed cube is $\chi_i = 4.8$, while it is $\chi_e = 24$ outside. A computational mesh of $49 \times 49 \times 25$ mesh points is used. The contour lines of the source function in the $x$-$z$ plane are shown.

The opacity outside the dashed cube is denoted with $\chi_e$. Inside the cube the opacity is reduced by a factor 5, and it is denoted with $\chi_i$. Here we use $\chi_e = 24$. Physically, this is a simplified model of a magnetic flux tube in an otherwise normal stellar atmosphere. In this tube the density and therefore the opacity are reduced.

For the radiative transfer calculations we assume radiative equilibrium and again the 'grey' case. From this it follows that the source function is equal to the mean intensity, i.e.

$$S = J.$$  \hspace{1cm} (7.16)
The mean intensity $J$ can be written as

$$J = \Lambda S + G$$  \hspace{1cm} (7.17)$$

where $\Lambda$ is the usual $\Lambda$-operator, and $G$ is the mean intensity due to the incident radiation. By combining the above equations, we get for the source function

$$S = \Lambda S + G.$$ \hspace{1cm} (7.18)

We show lines of constant $S$ in the $x$-$z$ plane in Fig. 35. Under the additional assumption of LTE, we have for the source function $S = B = \sigma T^4 / \pi$, where $T$ is the temperature. Therefore, the contour lines in Fig. 35 are also isotherms. As can be seen in the figure, the reduced opacity in the flux tube causes a cooling of the matter at the bottom and a heating at the top of the tube relative to the matter outside the flux tube (Kalkofen et al. 1989). This is qualitatively the same result as that of the 2D calculation of Steiner (1991).

In order to get the specific intensities in the $x$-$y$ layers within a reasonable number of steps, we rotate the set of vectors $\vec{n}$ so that no vector has a corresponding angle $\theta$ smaller than 8.5°. Then we need at most 25 steps to get the specific intensities of any $x$-$y$ layer in our example, but on average we have $\bar{N} = 12$. We summarize the convergence properties of different iteration schemes in Fig. 36 similarly to Fig. 32 for the Dirichlet boundary conditions. As an initial guess we have set the source function everywhere equal to the mean intensity at the bottom, where the radiation is approximately isotropic and the incoming radiation is given. As in the previous example, convergence is achieved in the fewest steps with the multi-grid method using three and four grids. But because of the relatively small number $\bar{N}$ here, the multi-grid
method as implemented here is not as efficient as some of the other convergence methods. Overall it is the orthomin method with a non-local operator which converges fastest. Using Eq. (7.10), it only takes nine iterations for the first step with the same accuracy as in the previous example. Again this number decreases rapidly. However, if memory space is insufficient, the orthomin method with a local operator works best.

7.6 Conclusion

We have developed a computer code to solve radiative transfer problems in three dimensions on the MasPar MP-1, which is a SIMD machine with 8192 processors. We adapt the short characteristic method (Kunasz & Auer 1988) for a 3D cartesian grid in order to solve the radiative transfer equation. To parallelize the computation, we map the cartesian grid onto the two-dimensional grid of the processors. Therefore, we are successful in parallelizing the radiative transfer calculations if the scaling of the computing time is independent of the number of grid points in the $x$ and $y$ axis, and if it is only linearly dependent on the number of grid points in the $z$ axis. This has to be seen in comparison to calculations on conventional scalar or vector machines, where the computing time scales as $N^3$ when there are $N$ grid points in each dimension. In contrast, it only scales as $N$ on this SIMD machine if the problem is parallelized completely. Because obtaining the solution of the radiative transfer equation is a highly non-local problem, it is not possible to achieve complete parallelization. However, in the case of Dirichlet boundary conditions, we can reduce the scaling of the computing time to $N^2$ in the worst case, and in a wide range of problems it reduces to $\bar{N} \times N$ where $\bar{N}$ is much
Figure 36. Maximum relative error of the source function after each iteration step. From top to bottom (at N = 20): classical A-iteration, ALI with local $\Lambda^*$, ALI with non-local $\Lambda^*$, multi-grid method with $n=2$, Ng acceleration, orthomin with local $\Lambda^*$, orthomin with non-local $\Lambda^*$, multi-grid method with $n=3$, multi-grid method with $n=4$. The orthomin acceleration overshoots slightly during the convergence. Therefore, the maximum relative error occurs at different points after different iteration steps, and it is not as smooth as other methods.
smaller than \( N \) and close to one. Even in the case of periodic boundary conditions, where potentially \( \bar{N} \gg N \), one has in many applications \( \bar{N} \ll N \). A SIMD machine is therefore very suitable for studying inhomogeneities in stellar atmospheres, where periodic boundary conditions become important.

In order to achieve the final source function in the NLTE case, we have to make iterations. We have tested various methods to accelerate the convergence. Most of these can be completely parallelized. Only the multigrid method as implemented here turns out not to be well suited for SIMD machines, because the timing depends only weakly on \( N \) in comparison to a conventional computer, and because many processors will be inactive during the computations. Overall, the orthomin acceleration in combination with the ALI method gives the smoothest convergence. As an approximate \( \Lambda \)-operator we test a purely diagonal matrix and a matrix constructed by including 26 nearest neighbors. As the inversion of this latter matrix is difficult, we solve the corresponding equation using the Jacobi iteration. While the diagonal matrix already gives a good solution in the effectively optically thick areas, it does not accelerate the convergence sufficiently fast in the optically thin areas, as expected. There the operator's including the 26 nearest neighbors gives a significant improvement. But because of the large memory space required to store this operator, it may not always be possible to use it.

Overall we find that it is possible to efficiently solve NLTE radiative transfer problems on a SIMD machine. We were careful to write the program without overly specializing for the particular SIMD machine that we have available. As the entire code is written in a computer language that is close to the
new FORTRAN 90 standard, we will even be able to adapt the code easily to conventional machines once this new standard is generally available. Furthermore, it should not be difficult to convert this program onto a different kind of SIMD machine, as we do not explicitly use specific hardware tools of the MasPar MP-1. The only way we have specialized the code for this particular machine is by assuming a 2D grid of processors, but this assumption should not prove to be very restrictive when using other SIMD machines.
CHAPTER 8

CONCLUSION

In this work we aimed to investigate various aspects of three-dimensional radiative transfer in astrophysical systems. Such calculations are very expensive to compute due to the large amount of memory and computing time they require, and therefore have not been possible until recent years. We developed methods to perform such calculations on a SIMD massively parallel computer. To our knowledge, radiative transfer calculations have not been made before on such a computer. We have used our codes to examine a variety of areas in astrophysics. These include the structure of neutral hydrogen disks of galaxies, the distribution of dust in the outer regions of galaxies, accretion columns onto magnetic white dwarfs, and general NLTE calculations for the two-level atom. Here we provide a brief overview of our results.

First we investigated the large scale structure of neutral hydrogen in disk galaxies. While the inner disks of such galaxies are essentially flat and axisymmetric, the outer regions of disk galaxies show strong warps (Bosma 1981). Barker et al. (1993) have derived an analytic description of this structure, and they have constructed model galaxies based on this formula. We then used our code to calculate the resulting column density and radial velocity profiles of neutral hydrogen for these model galaxies when seen from many different directions. We have thus produced a movie of a "fly-by" around our model
galaxies, and a few images of this movie are presented in this thesis. It hence turns out that the resulting profiles of one galaxy can vary strongly when one looks at it from different directions. The general properties of observed profiles of a number of galaxies could therefore be matched in this way.

A second astrophysical case that we investigated involved the distribution of dust in outer regions of galaxies. Normally one would not expect any dust far outside the optical disk of a galaxy because the production of dust is linked to the presence of stars. However, from our study of the 3D structure of galaxies we noticed that the HI disk of NGC 5055 starts to cross over the (inner) optical disk just where a radial asymmetry of the optical disk occurs as seen in the photograph of NGC 5055 by Sandage (1961). If dust were present in the outer HI regions, the observed radial asymmetry of this galaxy could arise naturally. Unfortunately, we could not determine unambiguously the amount of dust present because of the small-scale structure of this galaxy and because of the problem of determining the position angle of this galaxy accurately. However, we were able to reproduce the general qualitative characteristics of this galaxy by assuming the presence of dust in outer HI regions. If dust indeed exists far outside the (inner) optical disk, then this may have consequences for the star formation during a time when this galaxy was just forming.

Next we examined accretion columns onto magnetic white dwarfs. In these systems the matter emits strong cyclotron radiation and one has to solve the transfer equation for polarized radiation. In order to do this we first had to examine the numerical calculation of the corresponding opacities. After that
we computed the spectrum emitted from an AM Her system, which we compared to previously published calculations (Wu 1989). Then we investigated the radiation from extended polar caps (Chanmugam & Frank 1987). These authors calculated the spectra of such systems by regarding the emission region as a two-dimensional region and obtaining the length of the rays through the shock from trigonometric considerations. While this appears to be a sufficient approximation at first sight because of the relatively small shock height, it turns out that this approximation leads to a significant underestimate of the amount of emitted radiation. To calculate the spectra of such systems one has to consider the full three-dimensional extent of the shock and fully three-dimensional radiative transfer calculations are necessary. Finally, we modeled the radiation from the intermediate polar RE 0751+14, which was discovered just recently (Mason et al. 1992) and is only the second object of its class to have significant polarized radiation (Rosen et al. 1993; Piirola et al. 1993). By fitting the calculated spectrum to the observed one we find that this IP must have a magnetic field of ~ 10 MG. It has previously been suggested that IPs can evolve into AM Her systems (Chanmugam & Ray 1984; King et al. 1985), but such an evolution requires that IPs have strong magnetic fields (> 3 MG; Chanmugam & Ray 1984). Our findings therefore support such an evolutionary scenario.

Finally, we wrote a code on this massively parallel computer to solve NLTE problems. In NLTE the source function is also a function of the radiation field that it produces. Calculating the source function in NLTE is therefore
a non-linear problem. To solve it we adapted the so-called short characteristic method (Kunasz & Auer 1988) to this computer and calculated the source function iteratively. We investigated the convergence behavior of several iteration methods and also examined the timing of these methods on this SIMD machine. As one can expect, significant differences to a normal scalar machine occur. We tested the program on two problems. The first one involved a case with Dirichlet boundary conditions, i.e. when the incoming radiation is known. In a second problem we examined our code when periodic boundary conditions are given. These conditions occur for example when one examines the transfer of radiation in a 3D stellar atmosphere. Such calculations become necessary when the 3D structure of convection cells is examined. We found that the MasPar with its SIMD architecture is particularly well suited for this problem.

As we have shown above, we were able to investigate with our numerical codes several very different areas of astrophysics, and we obtained important results. However, we want to point out that applications for our numerical codes are by no means limited to the above examples. Instead we can easily adapt our radiative transfer codes – particularly the imaging tool – to investigate many other astrophysical problems.
REFERENCES


Barker, K., Tohline, J.E., Frank, J. 1993, BAAS, 25, 798


Bornatici, M., Cano, R., De Barbieri, O., Engelmann, F. 1983, Nucl. Fusion, 23, 1153


Cannon, C.J. 1973a, JQSRT, 13, 627


129
Frederickson, P.O., McBryan, O.A. 1988, in Multigrid Methods, (New York: Dekker), 195
Goldstein, H. 1980, Classical Mechanics (Reading: Addison-Wesley)
Hackbusch, W. 1985, Multi-grid methods and applications (Berlin: Springer)
Hockney, R.W., Jesshope, C.R. 1988, Parallel Computers 2 (Bristol Adam Hilger)
Jeffery (Berlin: Springer), 377


Klein, R.I., Castor, J.I., Greenbaum, A., Taylor, D., Dykema, P.G. 1989, JQSRT, 41, 199

Kunasz, P., Auer, L.H. 1988, JQSRT, 39, 67

Kunasz, P., Olson, G.L. 1988, JQSRT, 39, 1

Lang, K.R. 1986, Astrophysical Formulae (Berlin: Springer)


Mihalas, D. 1978, Stellar Atmospheres (San Francisco: Freeman)

Mihalas, D., Binney, J. 1981, Galactic Astronomy (San Francisco: Freeman)


Olson, G.L., Auer, L.H., Buchler, J.R. 1986, JQSRT, 35, 431

Olson, G.L., Kunasz, P.B. 1987, JQSRT, 38, 325


Rybicki, G. 1971, JQSRT, 11, 589


Inst. Washington)

Sanders, R.H. 1990, A&AR, 2, 1


Stenholm, L.G., Störzer, H., Wehrse, R. 1991, JQSRT, 45, 47

Stoer, J., Bulirsch, R. 1990, Numerische Mathematik 2 (Berlin: Springer)


Tamor, S. 1978, Nucl. Fusion, 18, 229

Tohline, J.E. 1983, in The Internal Kinematics and Dynamics of Galaxies, ed. E. Athanassoula (Dordrecht: Reidel), 205


Vinsome, P.K.W. 1976, in Proc. of the 4th Symp. on Reservoir Simulation, Soc. of Petroleum Engineers (Boulder), 149


APPENDIX A

POLARIZED RADIATIVE TRANSFER

Polarized radiation can be described by the Stokes parameters $I$, $Q$, $U$, $V$ (Rybicki & Lightman 1979). The transfer equation in the Stokes formalism is given in Eq. (6.1) and the Faraday mixing coefficients are given in Eqs. (6.2) and (6.3). In the following we will derive the solution to this system of linear differential equations under the condition that all the absorption coefficients are constant along the line of sight. We will largely follow the derivation given by Pacholczyk (1977) but correct for some printing errors.

The solution can be written as $X = X_P + X_H \exp(-\kappa l)$ where $X$ stands for any one of the Stokes parameters. The particular solution $X_P$ satisfies

$$
\begin{pmatrix}
\kappa & q & 0 & v \\
q & \kappa & f & 0 \\
0 & -f & \kappa & h \\
v & 0 & -h & \kappa
\end{pmatrix}
\begin{pmatrix}
I_P \\
Q_P \\
U_P \\
V_P
\end{pmatrix} =
\begin{pmatrix}
\epsilon_I \\
\epsilon_Q \\
\epsilon_U \\
\epsilon_V
\end{pmatrix}.
$$

(A.1)

We get as the solution

$$
I_P = M^{-1} \left\{ \epsilon_I \kappa (\kappa^2 + 2m) - \epsilon_Q (rh + \kappa^2 q) - \epsilon_V (rf + \kappa^2 v) \right\}
$$

(A.2)

$$
Q_P = M^{-1} \left\{ -\epsilon_I (rh + \kappa^2 q) + \epsilon_Q \kappa (\kappa^2 - v^2 + h^2) + \epsilon_V \kappa (vq + fh) \right\}
$$

(A.3)

$$
U_P = M^{-1} \left\{ -\epsilon_I \kappa p + \epsilon_Q (-rv + \kappa^2 f) + \epsilon_V (qr - \kappa^2 h) \right\}
$$

(A.4)

$$
V_P = M^{-1} \left\{ \epsilon_I (-fr - \kappa^2 v) + \epsilon_Q \kappa (qv + fh) + \epsilon_V \kappa (\kappa^2 + f^2 - q^2) \right\}
$$

(A.5)
where

\[ M = \kappa^4 + 2(m - n)\kappa^2 - r^2 \quad (A.6) \]
\[ m = (f^2 + h^2)/2 \quad (A.7) \]
\[ n = (q^2 + v^2)/2 \quad (A.8) \]
\[ p = qf - vh \quad (A.9) \]
\[ r = vf + qh. \quad (A.10) \]

In the special case of local thermal equilibrium we have \( \epsilon_I = \kappa B, \epsilon_Q = qB, \epsilon_V = vB \) where \( B \) is the Planck function. Then the above equations simplify to \( I_P = B, Q_P = U_P = V_P = 0. \)

The homogeneous solution \( X_H \) satisfies

\[ I'_H = -qQ_H - vV_H \quad (A.11) \]
\[ Q'_H = -qI_H - fU_H \quad (A.12) \]
\[ U'_H = fQ_H - hV_H \quad (A.13) \]
\[ V'_H = -vI_H + hV_H \quad (A.14) \]

where the prime stands for the spatial derivative. For \( X_H \) one can write (Pacholczyk (1977) had a misprint in the indices; the equation is 3.159)

\[ I_H = \sum K_i \exp(\alpha_i l) \quad (A.15) \]
\[ Q_H = -\frac{1}{r} \sum \alpha_i (va_i + h)K_i \exp(\alpha_i l) \quad (A.16) \]
\[ U_H = - \sum a_i K_i \exp(\alpha_i l) \quad (A.17) \]
\[ V_H = \left\{ \frac{1}{r} \sum \alpha_i (qa_i - f) K_i \exp(\alpha_i l). \right\} \quad (A.18) \]

By inserting these equations into Eqs. (A.11) to (A.14) one can see that the
the constants $a_i$ and $\alpha_i$ must fulfill the following equations

$$-(1/r)a_i^2(va_i + h) = -q + fa_i \quad (A.19)$$

$$ (1/r)a_i^2(qa_i - f) = -v - ha_i. \quad (A.20) $$

One then gets that

$$a_1 = a_2 = (m + n - R)/p \quad (A.21)$$

$$a_3 = a_4 = (m + n + R)/p \quad (A.22)$$

$$\alpha_1 = -\alpha_2 = (n - m + R)^{1/2} = \lambda \quad (A.23)$$

$$\alpha_3 = -\alpha_4 = i(m - n + R)^{1/2} = i\mu \quad (A.24)$$

where $R = ((m + n)^2 - p^2)^{1/2}$. The constants $K_i$ can be obtained from the boundary condition at $l = 0$. There the incoming radiation is $X_0 = X_P + X_H$.

We get

$$K_I + K_Q = \frac{-a_3(I_0 - I_P) - (U_0 - U_P)}{-a_3 + a_1} = A_1 \quad (A.25)$$

$$K_I - K_Q = \frac{r (qa_3 - f)(Q_0 - Q_P) + (va_3 + h)(V_0 - V_P)}{\lambda - (va_1 + h)(qa_3 - f) + (va_3 + h)(qa_1 - f)} = A_2 \quad (A.26)$$

$$K_U + K_V = \frac{(U_0 - U_P) + a_1(I_0 - I_P)}{-a_3 + a_1} = A_3 \quad (A.27)$$

$$K_U - K_V = \frac{1}{i\mu} \frac{r -(va_1 + h)(V_0 - V_P) - (qa_1 - f)(Q_0 - Q_P)}{(va_1 + h)(qa_3 - f) + (va_3 + h)(qa_1 - f)} = \frac{1}{i}A_4. \quad (A.28)$$

It follows that

$$K_I = \frac{1}{2}(A_1 + A_2) \quad K_U = \frac{1}{2} \left( A_3 + \frac{1}{i}A_4 \right) \quad (A.29)$$

$$K_Q = \frac{1}{2}(A_1 - A_2) \quad K_V = \frac{1}{2} \left( A_3 - \frac{1}{i}A_4 \right). \quad (A.30)$$

These equations are in a much simpler form than Eq. (3.162) given by Pacholczyk (1977). Now we get as the general solution of the transfer equation

$$I = I_P + \{ A_1 \cosh \lambda l + A_2 \sinh \lambda l + A_3 \cos \mu l + A_4 \sin \mu l \} e^{-\mu l} \quad (A.31)$$
\[ Q = Q_P + \left\{ \frac{-q + f a_1}{\lambda} [A_1 \sinh \lambda l + A_2 \cosh \lambda l] \right. \\
+ \left. \frac{-q + f a_3}{\mu} [A_3 \sin \mu l - A_4 \cos \mu l] \right\} e^{-\kappa l} \] (A.32)

\[ U = U_P + \left\{ -a_1 [A_1 \cosh \lambda l + A_2 \sinh \lambda l] \right. \\
- \left. a_3 [A_3 \cos \mu l + A_4 \sin \mu l] \right\} e^{-\kappa l} \] (A.33)

\[ V = V_P + \left\{ \frac{-v - h a_1}{\lambda} [A_1 \sinh \lambda l + A_2 \cosh \lambda l] \right. \\
+ \left. \frac{-v - h a_3}{\mu} [A_3 \sin \mu l - A_4 \cos \mu l] \right\} e^{-\kappa l}. \] (A.34)

Again this expression is much easier to check than the equivalent one given by Pacholczyk (1977), which is Eq. (3.163). Note that the definitions of \( a_1 \) and \( a_3 \) as given by Pacholczyk (1977) and Meggitt & Wickramasinghe (1982) are reversed.

So far we have only shown how to calculate the Stokes parameters along some line of sight along which the magnetic field is homogeneous and the field lines are in the \( y-z \) plane. In the following we describe a recipe to solve the transfer equation in the case of a non-homogeneous magnetic field. First we divide the line of sight along each ray, that propagates through the emission region, into small regions in which all absorption coefficients are approximately constant and the magnetic field is homogeneous. Let us assume that at the beginning of such a region we know the Stokes parameters and that the magnetic field in some local coordinate system is in the \( y'-z \) plane with the \( z \) axis in the direction of propagation. As all rays have the same direction vector \( n \) the \( z \) axis is the same for all rays and will remain fixed in space throughout the calculation. However, the \( y' \) axis will in general vary from ray to ray and along some ray it will vary from region to region. Now we solve the
transfer equation with the above formulae for the current region on some ray and obtain the Stokes parameters $I', Q', U', V'$ at the end of this region. Then we transform the primed Stokes parameters from the local frame into a global unprimed frame that is identical for all rays for the particular direction vector $n$. The $z$ axis of this global frame is parallel to $n$ and the $y$ axis is fixed for the particular direction vector $n$. In this global frame the Stokes parameters are unprimed and with an angle $\Psi'$ between the $y$-$z$ plane and the $y'$-$z$ plane we can write (Chanmugam & Frank 1987)

\[
Q' = Q \cos 2\Psi' + U \sin 2\Psi' \quad (A.35)
\]
\[
U' = -Q \sin 2\Psi' + U \cos 2\Psi' \quad (A.36)
\]

while $I' = I$ and $V' = V$. The angle $\Psi'$ can be calculated from

\[
\cos \Psi' = \frac{(n \times z) \cdot (B \times z)}{|n \times z| |B \times z|} \quad (A.37)
\]

where $z$ is the unity vector along some axis fixed in space for all viewing directions and $B$ is the local magnetic field vector. If in the next region the magnetic field lies in a plane $y''$-$z$ and forms an angle $\Psi''$ with the $y$-$z$ plane then we can transform the unprimed Stokes parameters into the new double primed system by substituting $\Psi'$ by $\Psi''$ and performing the inverse transformation as the one expressed in Eqs. (A.35) and (A.36). After that we can again solve the transfer equation as before but now in the double primed system. We repeat this procedure along all rays and end up with the Stokes parameters for all rays defined in the same global frame. Then we can add the Stokes parameters of the different rays and thus calculate the integrated spectrum of the emission region.
In the following we derive equations for the formal solution for some special cases. First we get the formal solution for the case when there is no incoming radiation and the matter is in local thermal equilibrium. In this case $A_2 = A_4 = 0$. By noting that $a_1 = 1/a_3$ and $-a_3 + a_1 = -2R/p$ we get

\begin{align*}
I/B &= 1 - (p/2R) (a_3 \cosh \lambda l - a_1 \cos \mu l) e^{-\kappa l} \quad \text{(A.38)} \\
Q/B &= -(p/2R) \left( \frac{f - qa_3}{\lambda} \sinh \lambda l - \frac{f - qa_1}{\mu} \sin \mu l \right) e^{-\kappa l} \quad \text{(A.39)} \\
U/B &= (p/2R) (\cosh \lambda l - \cos \mu l) e^{-\kappa l} \quad \text{(A.40)} \\
V/B &= -(p/2R) \left( \frac{h + va_3}{\lambda} \sinh \lambda l + \frac{h + va_1}{\mu} \sin \mu l \right) e^{-\kappa l} \quad \text{(A.41)}
\end{align*}

These are the same equations as those given by Meggitt & Wickramasinghe (1982) except that these authors have $a_1$ denoted as $a_3$ and vice versa.

Finally, we want to look at one further special case. In the case of cyclotron absorption one has for sufficiently large harmonic numbers (about $s > 5$)

\begin{equation}
q^2 + v^2 \ll h^2 + f^2. \quad \text{(A.42)}
\end{equation}

Then one can expand the above equations in terms of $n/m$. If one further has the condition that

\begin{equation}
(h^2 + f^2)^{1/2} l \gg 2\pi \quad \text{(A.43)}
\end{equation}

then the above equations reduce to

\begin{align*}
I/B &= 1 - \exp(-\kappa l) \cosh \lambda' l \quad \text{(A.44)} \\
Q/B &= \text{sgn}(r) \left[ \frac{h}{(f^2 + h^2)^{1/2}} \right] \exp(-\kappa l) \sinh \lambda' l \quad \text{(A.45)} \\
U/B &= 0 \quad \text{(A.46)} \\
V/B &= \text{sgn}(r) \left[ \frac{f}{(f^2 + h^2)^{1/2}} \right] \exp(-\kappa l) \sin \lambda' l. \quad \text{(A.47)}
\end{align*}
where $\lambda' = |v f + q h|/(f^2 + h^2)^{1/2}$ and is derived from $\lambda$ under the assumption that $n/m \ll 1$. One can see that the ratio of linear to circular polarization is

$$Q/V = h/f = \sin^2 \theta/(\omega/\omega_c) \cos \theta.$$  \hfill (A.48)

When the angle $\theta$ between the magnetic field and the line of sight is small then $h \approx 0$. The propagation is called quasi-longitudinal and one has the situation of large Faraday rotation. Using Eq. (6.2) we can write this condition as

$$\frac{\omega_p^2 \omega_c}{\omega^2 c} l \gg 2\pi.$$  \hfill (A.49)

When $\theta \approx \pi/2$ then $f \approx 0$ and the propagation is called quasi-transverse. One has the situation of large Faraday pulsation and using Eq. (6.3) we get

$$\frac{\omega_p^2 \omega_c^2}{2 \omega^3 c} l \gg 2\pi.$$  \hfill (A.50)

Meggitt & Wickramasinghe (1982) stated that Eqs. (A.44) to (A.47) follow by only assuming Eq. (A.42). This, however, is incorrect. If it were correct, then the condition of large Faraday rotation and pulsation would not depend on the path length. As a consequence the ratio $Q/V$ would always be given by Eq. (A.48) and would be independent of the path length. But this is not the case. We illustrate this in Figs. 37 and 38 which compare the calculated ratio $Q/V$ from Eqs. (A.39) and (A.41) to the theoretical value of $Q/V$ in the case of large Faraday rotation and pulsation. We choose values of temperature and angle $\theta$ for which Meggitt & Wickramasinghe (1982) presented $I$ and the degree of linear and circular polarization in form of figures. Therefore our Fig. 37 corresponds to Fig. 3 and our Fig. 38 corresponds to Fig. 4 in Meggitt & Wickramasinghe (1982). As one can see $Q/V$ is strongly dependent on the
path length. Only at large path lengths, i.e. when Eq. (A.43) is fulfilled, does \( Q/V \) obey Eq. (A.48) and only in this case do we have large Faraday rotation and pulsation. For the cases considered here \( Q/V \) agrees best with the value for large Faraday rotation and pulsation for \( kT = 20 \text{ keV} \), a harmonic number of 15 and a path length of \( l/l_0 = 10^9 \). For angles between \( 2^\circ \) and \( 88^\circ \) the calculated value of \( Q/V \) deviates from its theoretical value no more than 0.001 \%. This is also a test for the accuracy of our transfer code as it results in the correct value of the Stokes parameters in this important limiting case.

Under the assumption of large Faraday rotation the set of Eqs. (6.1) that describe the radiative transfer decouple into two equations for the ordinary and extraordinary mode. They are given in Eq. (6.13) and have the formal solutions (with no incoming radiation)

\[
I_\pm = (B/2) \left[1 - \exp(-\alpha_\pm l)\right].
\]  

(A.51)

The Stokes parameters are calculated from (Ramaty 1969)

\[
I = I_+ + I_-
\]  

(A.52)

\[
Q = I_+ \left(\frac{1 - a_+^2}{1 + a_+^2}\right) + I_- \left(\frac{1 - a_-^2}{1 + a_-^2}\right)
\]  

(A.53)

\[
U = 0
\]  

(A.54)

\[
V = I_+ \left(\frac{2a_+}{1 + a_+^2}\right) + I_- \left(\frac{2a_-}{1 + a_-^2}\right)
\]  

(A.55)

where the functions \( a_\pm \) are given in Eq. (6.12).

One can easily show that Eqs. (A.51) to (A.55) combine to give Eqs. (A.44) to (A.47) if

\[
\alpha_\pm = \kappa \mp \lambda' \kappa + \text{sgn}(r) \left\{ \mp \frac{h}{\sqrt{f^2 + h^2}} q - \frac{\pm f}{\sqrt{f^2 + h^2}} \right\}
\]
Figure 37. Calculated value (solid) and value in case of large Faraday rotation and Faraday pulsation (dashed) of the ratio \( Q/V \) for \( kT = 20 \text{ keV} \). The harmonic number \( \omega/\omega_c \) increases from the top to the bottom row while the length \( l/l_0 \) increases from the left to the right column. The length unit \( l_0 = \omega_c c/\omega_p^2 \).
Figure 38: Same as the previous figure but for $kT = 50$ keV.
The latter relation follows as \( v < 0 \) for \( \theta < \pi/2 \) and \( q < 0 \) for all \( \theta \) for cyclotron opacities. This is the same equation that we proved formally in Chapter 6.2.

From the above equation we can also see that

\[
\kappa = (1/2)(\alpha_+ + \alpha_-). \tag{A.57}
\]

Finally, we want to point out that the transfer equation (6.13) in the ordinary and extraordinary mode and its solution expressed in Eqs. (6.14) to (6.18) cannot be used to solve the transfer equation in the general case of a non-homogeneous magnetic field along the line of sight. Even if we divide the line of sight into small homogeneous regions and if we assume that the condition of large Faraday rotation and pulsation within each region is fulfilled, we generally cannot transform to a new set of Stokes parameters with \( U = 0 \) as required by Eq. (6.17). The direction of the local magnetic field defines the local y-z plane (otherwise the emissivity corresponding to the Stokes parameter \( U \) is not zero). But in this coordinate system we will normally have \( U \neq 0 \). Therefore we have to use the transfer equation in the general Stokes formalism, i.e. Eq. (6.1) and its solution, in the case of magnetic fields that vary along the line of sight.
APPENDIX B

LETTER OF PERMISSION

Observatoire de Marseille
2 Place Le Verrier
13248 Marseille Cedex 4
FRANCE
Tel: (33) 91 95 90 86
Fax: (33) 61 62 11 90

DESTINATAIRE: Horst VATH
ADRESSE: Dept. of Physics, LSU, Baton Rouge, LA 70803
TELE-FAX: 19-1 604.385.8855
MESSAGE EXPEDIE PAR: Albert Bosma
DATE: June 7, 1984
COMMENTAIRES:

Dear Mr. Vath,


Yours sincerely,

Albert Bosma

NOMBRE TOTAL DE PAGES DU DOCUMENT: 1
VITA

Horst M. Väth was born on September 1, 1967, in Mannheim, Germany. He started his studies of physics in 1986 at the Ruprecht-Karls Universität in Heidelberg, Germany, where he received his Vordiplom in physics in July 1988. In 1989 he went to Louisiana State University in Baton Rouge, USA. Initially he worked on radial oscillations of neutron stars for which he received his M.S. in December 1990. After that he took his General Examination in April 1991 and then started the research for his dissertation. This research involved large scale computing of three-dimensional radiative transfer as it occurs in astrophysical systems. He finished this work in summer of 1994.
Candidate: Horst M. Väth

Major Field: Physics

Title of Dissertation:
Three-Dimensional Radiative Transfer on a Massively Parallel Computer

Approved:

Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

Date of Examination: 5/19/94